DOCUMENT RESUME

ED 036 902 EA 002 794

TITLE	American Educational Research Association Abstracts of Papers, Fiftieth Anniversary Meeting 1966.
INSTITUTION	American Educational Research Association, Washington, D.C.
PUB DATE	66
NOTE	135p.; Abstracts of papers presented at the Anniversary Meeting of the AERA (50th, Chicago, 1966)
EDRS PRICE	EDRS Price MF-$0.75 HC Not Available from EDRS.
DESCRIPTORS	*Abstracts, *Conference Reports, Counseling, Curriculum Development, Educational Environment, Educationally Disadvantaged, Educational Objectives, *Educational Research, Individualized Instruction, Measurement Techniques, *National Organizations, Programed Instruction, Research Methodology, Statistical Analysis, Student Attitudes, Student Motivation, Teacher Behavior, Test Construction

ABSTRACT

This document contains abstracts of 53 papers presented at the 1966 annual meeting of the American Educational Research Association. A sampling of the varied subjects covered includes: teacher training, behavior, and evaluation; student aptitudes, attitudes, motivation, and achievement; education of disadvantaged; statistical methods and theory; measurement problems; educational environments; curriculum development; test design; programed instruction; individualized instruction; language, mathematics and science learning; concept learning; and uses of computers in education. An index of the participants concludes the collection. Related documents are EA 002 791, EA 002 792, and EA 002 793. [Not available in hard copy due to fine print of original document] (DE)

ED036902

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

Abstracts of Papers

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Fiftieth Anniversary Meeting

Chicago 1966

EA 002 794

American Educational Research Association
1201 Sixteenth Street, NW
Washington, D.C. 20036

TABLE OF CONTENTS

Session 2--Perceptual and neurological factors in reading page 1
Session 4--Characteristics of teachers 3
Session 7--Administration and the educational program 6
Session 11--Individually prescribed instruction 9
Session 13--Educational programs for disadvantaged children 12
Session 14--Studies of university students' attitudes 14
Session 17--New directions in school-board research 17
Session 19--Language composition 19
Session 22--Mathematics and science learning 22
Session 23--Motivation and learning 25
Session 24--Counseling and counselor training 28
Session 26--Research and the teaching of reading 31
Session 31--Linguistic factors in verbal behavior 33
Session 32--Personal variables and the teaching process 35
Session 34--Curriculum development 37
Session 35--Research methodology 39
Session 38--Modifying teacher behavior 41
Session 41--Concept learning 43
Session 46--Instructional variables and curriculum 46
Session 47--Computers in measurement and research 48
Session 48--Studies of student characteristics II 50
Session 49--Statistical method and theory I 53
Session 52--Research on test-taking behavior 56
Session 54--Problem solving and creativity 58
Session 55--Teacher behavior and pupil growth 61
Session 57--Studies of student characteristics I 63
Session 58--Test development 66
Session 60--Modifying teacher behavior 68
Session 61--The environment for learning 71
Session 62--The validation of educational measures 73
Session 63--The school and the larger society 75
Session 64--Organizational variables: innovation and evaluation 77
Session 65--Student attitudes and achievement 79
Session 66--Speech and language learning in children 81
Session 67--Assessment of educational environments 83
Session 68--Intellectual development and its measurement 85
Session 71--Measurement of classroom process variables 87
Session 72--The culturally disadvantaged student 89
Session 73--Measurement in curriculum and instruction 91
Session 77--Statistical method and theory II 93
Session 78--Motivational factors in reading 95

TABLE OF CONTENTS (CONT'D.)

Session 79--Verbal learning 97
Session 82--Specific teacher behavior 100
Session 84--Programmed instruction 103
Session 85--Perceptual and motor skill learning 106
Session 87--Meaningful learning and transfer 108
Session 88--Measurement in language arts and reading 111
Session 90--Teacher training and evaluation 114
Session 91--Instructional objectives 116
Session 92--Factors in school achievement 118
Session 93--Research on teaching 121
Session 96--Characteristics of socially disadvantaged children 124
Session 98--Studies of student aptitude and achievement in universities 127
Index of names-- 129

INTRODUCTORY NOTE

At the 1965 meeting, brief abstracts of papers were available. This first attempt to provide participants with an outline of the proceedings was sufficiently successful that the decision was made to undertake a more ambitious publication venture for the 1966 meeting.

The abstracts were increased in length from 100 words to 250 words by the program committee in order to facilitate the review of each contribution. This immediately presented many problems as soon as the task of preparing the abstracts for the printer was started. The main problem was that few of them were limited to 250 words and many extended to 5,000 words. As soon as the first batches of abstracts were received by the program chairman, he sent out letters requesting briefer abstracts within the 250-word limit, for the cost of reproducing the material received would have been prohibitive. Those who replied within the time allowed provided briefer abstracts but many were still far too long to reproduce. For example, the participant who had submitted a 5,000 word abstract reduced his to 512 words. Few were below 300 words. Wordy people hate to discard words. By this time, the printer's deadline was approaching and copy had to be prepared. The decision had to be made to cut some of the abstracts. In many cases this was merely a matter of cutting the last paragraph which described future research. In some, it involved the removal of redundancy. I hope that those whose abstracts have been so treated will understand the problems of an editor who has a deadline to meet.

Space did not permit the reproduction of symposium abstracts. An exception to this is Session 11which is a symposium produced by the assembly of a number of individually submitted papers.

Robert M.W. Travers
Chairman, Program Committee
Western Michigan University

SESSION 2

Paper session--PERCEPTUAL AND NEUROLOGICAL FACTORS IN READING.

Arthur V. Olson, University of Georgia. The Relationship of the Frostig Developmental Test of Visual Perception to Achievement Test Scores and Specific Reading Abilities in Grade II.

The purpose of this study was to determine (1) the predictive value of the Frostig Developmental Test of Visual Perception to general achievement in the second grade and (2) the relationship between the Frostig Test and specific reading abilities, i.e. paragraph comprehension, word recognition, hearing sounds in words, visual memory, using reversible words in context correctly, and synthesizing words in context.

The population was composed of 29 girls and 42 boys in the second grade. The mean mental age of the group was 95.46 with a S. D. of 11.46; chronological age 89.20 months, S. D. 7.41, and a mean I.Q. of 104.35, S. D. 16.17.

Intercorrelations of the Frostig Test (total score) with the CAT shows that all correlations were significant at the .01 level. All subtests of the Frostig contributed to the correlation except for Test 3 (Form Constancy). The CAT arithmetic reasoning and fundamentals showed the highest correlation with the Frostig, .509 and .566 respectively.

The Frostig Developmental Test of Visual Perception was of little value in predicting the specific reading abilities of the students tested.

Joseph B. Tremonti, Loyola University. The Use of Neurological Organization Accompanying Developmental Reading Techniques for College Age Students with Reading Disability.

One of the most promising areas of research dealing with learning, especially reading, is that which concentrates on neurological organization. Since Orton, there have been those who stressed neurological factors and lateral dominance as etiology for poor learning and reading disability. Also, there have been those who were skeptical of the neurological approach, especially as it pertains to lateral dominance. Some studies have found no etiological relationship of dominance to reading disability, but have employed questionable procedures.

A recent pilot study of the neurological aspects related to reading and IQ scores was conducted by President Robert Morris and Reverend Joseph B. Tremonti, C.S.V. of the University of Plano. The concepts of neurological organization described by Doman, Delacato, et al., at The Institutes for the Achievement of Human Potential in Philadelphia, Pennsylvania were employed along with developmental reading techniques.

The study attempted to provide answers to the following questions: (1) What is the relationship, if any, between neurological organization and reading scores; and (2) What is the relationship, if any, between neurological organization and IQ scores? The experimental group consisted of thirteen students, ranging in age from nineteen to twenty-eight. Intelligence tests, reading tests, and tests using the reading-eye camera were administered at the beginning of the program of neurological organization and at the end. The observed differences in test scores were statistically significant. The findings indicated that a program of neurological organization may be helpful to college age students with reading disability.

SESSION 2

Shirley C. Feldmann, City University of New York; Dorothy E. Schmidt, and Cynthia P. Deutsch, New York Medical College. *A Study of the Effects of Auditory Training on Remedial Reading.*

Previous research shows a relationship between deficits in auditory perceptual skills and reading retardation, especially among children of low socio-economic status. This study examined the relation of auditory perceptual skills to reading skills by measuring the effectiveness of an auditory training program in facilitating the retraining of retarded readers.

Fifty-eight eight-year-old retarded readers of low socio-economic status were assigned to small groups of three to four children in which they received training in reading, auditory skills, both or neither (the control group). The developmental auditory program was specifically devised for the study. Batteries of auditory and reading tests were given to the subjects before and after the 5 month treatment period. The tests were administered two more times, once after a six-months interval and again after a twelve-months interval, in order to ascertain any long-term modifications in reading and auditory skills and in their possible interrelationships.

Analyses of covariance indicated that some of the reading and auditory tests showed treatment differences. Most of the tests showed differences due to time. i.e. development. These tests also showed some race and tutor differences. There were no strong indications that any particular group improved more than others on all of the reading and auditory tests.

John R. Bergan, University of Kansas and Gloria Macchiavello, University of Kansas. *Visual Imagery and Reading Achievement.*

This study investigates the relationship between visual imagery measured on a five point self-rating scale and reading comprehension measured by the Stanford Achievement Test. Twenty-four boys , 27 girls, from two 4th grade classes participated in the study.

A relationship between imagery and reading has often been suggested, but substantiating evidence has been lacking largely because of the inadequacy of imagery measures. Questionnaires which provide a practical starting point for measuring imagery are subject to response bias. In this investigation, bias was removed by the following procedure: The questionnaire was divided into three subtests, visual, auditory and kinesthetic. It was assumed that insofar as individuals rated themselves higher on some items than on others, inter-individual differences on the total test would tend to cancel. The degree to which they did not was taken as a measure of response bias. Standard scores based on each individual's ratings were used to equate total test results. Inter-individual differences on subtests were still present and reflected differences in which response bias could be assumed to be eliminated. Visual imagery, then, was defined as the total of an individual's standard scores on the visual imagery subtest. A correlation of .44 ($p = .01$) was obtained between the measure of visual imagery and reading achievement.

SESSION 4

Paper session --CHARACTERISTICS OF TEACHERS

Albert H. Yee, University of Texas. Factors Involved in Determining the Relationships Between Teachers' and Pupils' Interpersonal Attitudes.

In 1964-65, an empirical study investigated problems dealing with causality in the relationships between teacher-pupil attitudes. Pre-and post-test attitude measures of 100 teachers in grades 4, 5, and 6 and their c. 3,000 pupils from middle class neighborhoods were obtained. Teachers' attitudes were measured by the Minnesota Teacher Attitude Inventory (MTAI) and a specially prepared semantic differential. An inventory called "About My Teacher" measured pupils' attitudes and provided 15 measures on the dimensions of affective, cognitive, disciplinary, innovative, and motivational attitudes.

Cross-lagged panel correlation and a method of using chi square tests of resolutions in attitude relationships were applied to test hypotheses. Significant results showed this study's teachers caused pupils' attitudes to change more than pupils caused teachers' attitudes to change.

Results to be presented deal with the influence of teaching experience on teachers' relationships with pupils, and also relationships of teacher-pupil attitudes as determined by other variables, principals' MTAI scores and teachers' sex, factor analyses of attitude responses, and a new validation of the MTAI with pupils' and principals' ratings of teachers.

Philip R. Merrifield, O. L. Davis, Jr., and James A. Phillips, Jr., Kent State University. Factors in OScAR Ratings of Secondary Level Student Teachers.

OScAR Ratings were made of classroom behaviors of 115 student teachers and their pupils on three occasions during Fall, 1964. These ratings cover areas of teacher and pupil verbal and non-verbal activities. The prescribed methods of observation were used by supervising faculty members.

Eight of the 13 scores suggested by Medley and Mitzel were selected for analysis, based on the findings at secondary level reported by Bowers, Davis, and Bowers. The 24 variables -- eight scales on each of three occasions -- were intercorrelated; because many of the ratings had extremely skewed distributions, each of the 24 distributions were dichotomized as near its median as possible. Thus the intercorrelations were phi coefficients.

Of the principal components extracted from the intercorrelation matrix, six were rotated graphically to orthogonal simple structure. Five of the resulting factors may be named as follows: verbal supportive; verbal non-supportive; verbal problem-structuring; affection; and seatwork. The sixth factor is a residual. . The stability across occasions of these five ways of differentiating teacher behavior is high.

Trends across observation times, based on original data, and analyses of scales within observations will be reported. Implications of the findings for teacher training will be suggested.

SESSION 4

Richard L. Turner, Indiana University. *Beginning Teacher Characteristics and Beginning Teacher Problems: Some Predictive Relationships.*

For two successive years, all beginning, inexperienced teachers grades 1 - 6 in 13 Indiana school systems were assessed on two devices as they began teaching: 1. The Teacher Characteristics Schedule (TCS). 2. Teaching Tasks (TT) in arithmetic and reading. At the end of two years for the first group and one year for the second group, supervisors and principals reported to open end questions the types of problems experienced by these teachers. From these reports, one group experiencing no problems and six groups with one or more of the following problems in teaching were established: discipline, management, reading, subject matter, social-emotional and pupil expectancy.

Comparisons, via analysis of variance, of the TCS and TT scores of the "no problems" group with each problem group except the "social-emotional" group, which contained too few cases, indicated that the "no problems" group differed significantly (from $p<05$ to $p<.001$) from the other problem groups in a number of respects.

Conclusions: Certain beginning teacher problems are associated with antecedently measured teacher characteristics in a predictable way.

Gene Glass, University of Illinois. *Discovery of Factors in Teachers' Perception of Students.*

A partial answer was sought to the question: To what aspects of a student does a teacher attend when judging similarity to other students?

Descriptions were written of eight fictitious high school students, each of whom was at one or the other of two levels of *three* attributes: Aptitude (high-low), Home Environment (good-bad), Classroom Behavior (good-bad). For example, one "student" was very bright; behaved well in class; and came from a broken, poor home. The 8 descriptions of students were given to 24 teachers to judge the similarity of students to each other by the method of multidimensional rank order. The "students" were scaled in an eight-dimensional space into which were placed two factors which accounted for almost all of the inter-stimulus distances.

Interpretations of the factors left no doubt that Home Environment was ignored by the teachers. The interpretation of the two bi-polar factors present in the teachers' perceptions of the "students" was slightly equivocal. On the basis of distinctness of simple structure and feedback from the teachers after judging, the preferred interpretation of Factor A was that it represented congruence of Aptitude and Classroom Behavior: Factor A had large positive loadings from the two students who were bright and well behaved poorly in class. Factor B represented incongruence of Aptitude and Classroom Behavior.

SESSION 4

William J. Gnagey, Illinois State University. The Relationship of Parental Acceptance to the Professional Attitudes and Academic Achievement of Students in Teacher Education.

The primary purpose of this investigation was to ascertain the relationship between a student's perception of his own parental acceptance and: (1) his manifest anxiety, (2) his tendency to over- or under-achieve, (3) his attitudes about teaching public school students, (4) his tendency to shift these attitudes toward an instructor's position, (5) his attitudes toward the professional education curriculum, (6) his tendency to be authoritarian, and (7) his tendency to hold achievement-oriented or affiliation-oriented beliefs.

Four representative Educational Psychology classes (N= 139)were drawn the fall semester offering at a large Midwestern teachers college. Two additional classes (N=53) were selected from a small Eastern liberal arts school also training teachers.

At the beginning of the semester, all subjects were administered measures of parental acceptance, manifest anxiety, academic ability, academic achievement, attitudes toward teaching, attitudes about professional education, authoritarianism, and achievement- or affiliation-oriented beliefs. At the end of the same semester, subjects completed a student information sheet and a second measure of attitudes toward teaching.

Results. Significant correlations were obtained between parental acceptance scores and measures of anxiety, attitudes toward teaching, and attitudes toward professional education. A significant difference was found between the parental acceptance scores of male under-achievers and overachievers. No significant findings appeared suggesting a relationship between parental acceptance and attitude shift, authoritarianism, or tendency to hold achievement - or affiliation-oriented beliefs.

SESSION 7

Paper session-ADMINISTRATION AND THE EDUCATIONAL PROGRAM

Bryce M. Fogarty, University Council for Educational Administration
Characteristics of Superintendents of Schools and Centralization.

The focus of this study was on the relationships between the degree to which decision making is centralized in the role structures of school systems and selected characteristics of superintendents of schools. Leader behavior was also investigated in relation to personality factors.

The following relationships were examined in this study: 1) superintendent's personality factors and degree of centralization of decision-making; 2) superintendent's personality factors and leader behavior; and 3) leader behavior and degree of centralization of decision making.

More than 2400 professional staff members in twenty school systems responded to the Decision Point Analysis, which provided data for computing the degree of centralization of decision making. The responses of the superintendent of each system in Cattell's 16 P. F. Questionnaire and to the Leader Behavior Description Questionnaire were obtained.

The instrumentation and methodology permitted determination of the perceived degree to which decision making was centralized in functional areas of administration and in school systems. No significant relationship between personality factors of superintendents and either their leader behavior or centralization of decision making was established; however, school systems in which decision making was centralized were administered by superintendents who tended to be intelligent, self-sufficient, tense, aloof, silent, practical and self-confident. A statistically significant relationship between leader behavior and centralization of decision making was not established although high scores by superintendents on the leader behavior dimension of Consideration tended to be associated with decentralization of decision making.

Donald C. Francke, Blue Island, Illinois, Community High School District 218. Personal Variables Related to Perceptions of Decision Making Responsibilities.

In order to understand better the organizational behavior of teachers, the relationships between certain personality and biographical variables and their perceptions of the locus of decision-making responsibilities, in addition to other potentially meaningful organizational variables, were examined.

In order to elicit the primary data, a Decision Point Analysis, several psychometric instruments, and a biographical questionnaire were administered to more than 6,000 teachers and administrators in the population of 31 Wisconsin school systems having 100 to 700 professional staff. Structured interviews were conducted, and measures of leader behavior, curriculum planning and implementation, and certain school system variables were assessed for sampled teachers and school systems.

Analysis of the data, primarily through correlation and regression techniques, revealed numerous relationships significant above the five percent level of confidence, with many of these above the one percent level. The most meaningful of these are: 1) Teachers' personality and biographical characteristics and their perceptions of decision-making responsibilities.

2) Certain administrative factors and curriculum characteristics of school systems; 3) Teachers' perceptions of their decision-making responsibilities and those of their administrators; 4) The leader behavior of administrators as perceived by themselves and their teachers, and curriculum development characteristics.

Eugene W. Tornow, Massewa, New York Central Schools. Administrators' Behavior and Teachers' Perceptions of Decision Points.

This study dealt primarily with the relationship of administrators' behavior and teachers' perceptions of decision points in a sample of the tasks involved in the operation of local school systems. Focus was upon the relationships between the congruence of teachers' perceptions of decision points and the administrative interrelationships of the superintendent of schools, the director of instruction, and the high school principal in eleven local public school systems.

Congruence of staff perceptions of decision points for each school system was established from the data obtained in the U.S.O.E. Project 1913. The study of the behavior of administrators was limited to certain personal variables, assessed through the administration of the California Psychological Inventory, and the interaction of the members of each administrative triad, ascertained through the administration of a structured interview.

Conclusions centered upon the relationship, or lack of a relationship between 1) congruence among teachers' perceptions of decision points within local schools and the nature of administrators' interaction; 2) congruence in the personal variables of administrators and the nature of administrators' interaction; and 3) the degree of congruence in the personal variables of administrators and the degree of congruence among teachers' perceptions of decision points.

Emmet J. Duffy, Duluth, Minnesota Central High School. The Role of the Director of Instruction.

Four directors of instruction in school systems were observed in terms of the tasks, interactions, and processes which characterized their roles. A comparison was made between the observed roles and the directors' perceptions of their roles. The extent to which the observed tasks are similar to those suggested by the 25 decision items included in the Decision Point Analysis instrument also was examined.

Each director was observed for three weeks. The frequency of the behaviors and the amount of time utilized were determined for each of the specific tasks and interactions. These data were compared with those obtained from the Decision Point Analysis instrument.

A director of instruction devotes approximately 70 percent of his observed time and frequency of behaviors to the tasks of curriculum and instruction and of staff personnel. Approximately 80 percent of his time is spent in interactions with people. Face-to-face verbal communication characterized over 65 percent of all interactions.

The two administrative task areas which receive the highest proportions of observed time and of frequency of behaviors of the directors of instruction are also the ones which the largest proportion of the professional staffs, and the directors themselves, perceive the director of instruction as being involved to the greatest extent in the decision-making process.

The observational data concerning the behavior of the director of instruction indicate that 15 of the 25 Decision Point Analysis items were suggestive of the tasks of the director of instruction.

Charles E. Kline, Purdue University. *Leader Behavior, Curriculum Implementation and Curriculum Change*.

The relationship of certain leader behaviors of the central office staff member credited with curriculum responsibility and curriculum plan implementation and curriculum plan change formed the focal points of a perceptual study conducted in ten Wisconsin public school systems. The central office curriculum decision maker of each system was identified through administering the Decision Point Analysis Instrument by U. S. O. E. Project #1913 staff. Curriculum plans spanning grades K to 12 were identified and sampled in each system according to determined criteria; similarly, 325 teachers denoted as instructors in curriculum areas included by the forementioned plans were selected, proportionate in number per system to the teacher populations of the ten systems.

The central office curriculum decision maker and selected teachers of each school system responded in terms of the central office curriculum leader to a Leader Behavior Description Questionnaire instrument, through which two behavioral dimensions were identified - initiating Structure and Consideration. In individual interviews, the subject's responses were recorded on a Curriculum Implementation index instrument which contained two scales - one measuring curriculum plan implementation and the other, extent of change in those plans.

Examination of correlations suggest among other findings that teachers' perceptions of central office curriculum leaders' behaviors indicative of *Consideration* are significantly correlated (.05) to teachers' perceptions of *implementation* of curriculum plans. Agreement between central office curricular leaders and teachers on leaders' behaviors indicative of *Initiating Structure* are negatively correlated (.01) to teachers' perceptions of implementation of curriculum plans. There is no significant relationship between perceived leader behavior, nor agreement on perceived leader behavior, and extent of curricular change.

SESSION 11

SYMPOSIUM --INDIVIDUALLY PRESCRIBED INSTRUCTION

Robert Glaser, University of Pittsburgh. The Program for Individually Prescribed Instruction

A major project of the University of Pittsburgh's Learning Research and Development Center during its first year of operation has been its program for Individually Prescribed Instruction (IPI) in which it has used as its field laboratory the Oakleaf School of the Baldwin-Whitehall school district. The project is a development project designed to produce a workable model for making certain provisions for individualizing instruction in the elementary school.

The work to date has been centered upon the development of curricula, teaching materials, and evaluation instruments to permit pupils to progress in reading, arithmetic, and science at a pace suitable to individual abilities and background. Specific instructional procedures are being developed to implement this program which require a restructuring of the functions of school personnel. Analyses of pilot-study data indicate that a significant degree of individualization can be achieved.

Additional work involves the development and adaptation of a variety of types of materials that will result in further individualization and will emphasize student self-direction and self-evaluation. One area for a concentration of effort is the construction of curriculum-embedded tests and the development of content-references scoring procedures for proficiency tests.

This paper describes the results from the first year of operation of the IPI program.

C. M. Lindvall and John Yeager, University of Pittsburgh. Correlates of Rate of Learning Under Individually Prescribed Instruction.

Investigations of the correlates of rate of Learning have been carried out as laboratory experiments by a number of psychologists and educators. However, studies of this important variable as it operates in actual classroom situations have been very limited in number due largely to the relative scarcity of classroom procedures which permit variability in individual student rates.

The program for Individually Prescribed Instruction permits pupils to work at individual paces in progressing through a sequence of units in each of three subject areas; reading mathematics, and science. In the research associated with this development program, studies of rate of learning are of central interest. Analyses of data obtained during the first year of the project and based on 159 elementary school students indicate that: 1) Within a given subject (e. g., arithmetic) in which a student spends varying amounts of time in different sub-areas or units, a reliable measure of overall rate of learning can be obtained. That is, there are significant differences among students over units; 2) There is no significant correlation between score on a test of scholastic aptitude and rate of learning.

The study also provides detailed information concerning correlations among rates in different subjects and units and between these sub-measure and different measures of aptitude.

John O. Bolvin, University of Pittsburgh. Variability of Pupil Achievement Under Individualized Instruction.

There is considerable evidence to support the assumption that pupils in a given grade achieve at varying levels in the same subject. One hypothesis for which supporting evidence is available is that the variability in a particular grade approximates the number of years the pupils have been in school (e.g. in the 3rd grade, one would expect a spread of 3 years in achievement). These results are reported for graded schools in which the materials, textbooks, school structure and pupils are graded. Educators, taking into account individual differences in achievement at a given grade level, propose and implement programs of individualized instruction where, among other innovations, grade levels are abolished. One of these, the Oakleaf Individually Prescribed Instruction project, provides the data for this study.

It is hypothesized that pupils involved in the individualized program will exhibit greater variability than pupils involved in a graded program. Data concerning pupil variability in achievement prior to and after one year involvement in the individualized program has been collected. Preliminary analysis suggests the confirmation of the hypothesis.

Richard Cox, University of Pittsburgh. The Development of a Sequentially Scaled Achievement Test.

In a program of individualized instruction,achievement tests are designed to discriminate between subjects who have mastered certain behaviors and those who have not. When the test is comprised of items which measure a variety of behaviors, it becomes desirable to have a convenient way to determine which behaviors the subject has or has not mastered without looking at his response to each item. This situation calls for an achievement test in which the score would indicate the response pattern of the subject. Such cumulative scales have been developed for the measurement of attitudes; attempts to apply scaling technique such as "scalogram analysis" to achievement tests have generally been fruitless.

A list of elementary math objectives were arranged in a postulated sequential order. Items designed to assess the behavioral objectives were written and combined into test form. This test was administered to a tryout group and the scalogram analysis was employed. This procedure indicated that a partial rearrangement of objectives and subsequently test items was necessary if the response patterns were to yield scale types. The resultant coefficient of reproducibility was .93, an acceptable criterion for validity.

A validation study was conducted utilizing the revisions suggested by the initial analysis. Again the data suggest that it is actually reasonable to develop a sequentially scaled achievement test.

Joseph I. Lipson, University of Pittsburgh. Transfer of Attained Skills to New Skills in Mathematics.

An important aspect of a program which permits pupils to progress through a curriculum sequence at individual rates is the extent to which pupils can demonstrate mastery of a given unit in the sequence without actual study in the units. This phenomenon is associated with the ability of a student to extrapolate from what he learns in one unit to the objectives that he is expected to master in a subsequent unit.

The program for Individually Prescribed Instruction permits an investigation of this ability in that the units of study are arranged in a carefully sequenced order and all students take a pre-test and a post-test on each unit. Data have been gathered to determine characteristics of students who display this ability to "pre-test out" of a given unit and to investigate the special characteristics of units which are unusual in the number of students who demonstrate mastery of the content on the basis of pre-test performance.

This paper presents data on the correlation of this ability to extrapolate one's knowledge and such characteristics as academic aptitude, achievement, and the rate of progression. It also presents the rationale for the use of data concerning the special characteristics of units that have a high proportion of pupils "pre-testing out" in the development of units of study which enhance the probability of the occurrence of this phenomenon.

SESSION 13

Paper session --EDUCATIONAL PROGRAMS FOR DISADVANTAGED CHILDREN

Sister Josephina, Boston College. An Academic Evaluation of a Head Start Program.

The "topsy" like growth of Head Start Programs demanded on the spot planning for the eight weeks program utilizing effective techniques for 196 all-white subjects (Charlestown, Mass.). To remedy obvious handicaps, the writer and teachers (N=13) drew up a curriculum emphasizing language competencies accompanied by pre-school activities. A daily log for each child, the teacher's plan and a daily schedule of activities supplied discussion data at the weekly conference. Visitation of classes gave opportunity for the teacher to present problems. Individual and group conferences followed, at which time discussions took place concerned with child growth, analysis of behavior patterns, sociology of the urban child, techniques of teaching and assessing readiness emphasizing the sensorial approach (Montessori), and use of research findings (Bruner, Piaget, Durkin, Suppes).

To assess growth in terms of mental ability and individual learning resulting from the semi-structured daily lessons, data from pre and post-readiness tests (Metropolitan-Form R & S) and a non-verbal intelligence test (Pintner-Cunningham) were statistically analyzed. There were devised individual cumulative folders containing a master card for test results, anecdotal records, results of interviews, daily logs, weekly accomplishment and pupil's work. Not only pupil interest and zest for learning evolved, but also sincere expressions of parental satisfaction and teacher amazement with the results of the Head Start Program. A varied semi-structured curriculum with major emphasis on the total growth of each child achieved measurable and tangible results.

E. Kuno Beller, Temple University. The Impact of Pre-School Experience on Intellectual Development in Educationally Deprived Children.

The present study investigated the impact of one year of pre-school education on the child's intellectual performance, on his attitudes toward the test situation, and on the child's ability to learn from his own experience.

The subjects were 100 children ranging in age from 46 to 76 months and attending nursery and kindergarten in four different schools. All children came from urban, lower class, educationally deprived educationally deprived neighborhoods; and the majority from Negro families and broken homes. Children with and without prior nursery experience were compared: 1) on their intelligence scores derived from 4 different tests, i.e. the Stanford Binet Test, the Peabody Picture Vocabulary Test, the Philadelphia Verbal Abilities Test, and the Goodenough Draw-A-Man Test; 2) on 8 scales that were constructed to measure the children's attitudes toward the test situation, e.g., involvement, persistance, inhibition, and rigidity; 3) on their learning a problem-solving task in 30 trials under 2 different conditions, i.e., extrinsic and intrinsic reinforcement.

Results: 1) Prior nursery experience resulted in superior intellectual performance, which emerged *most* consistently on tests involving manipulations of verbal cues, particularly on the Stanford Binet Test, and least consistently on the Goodenough Test, a performance test that involved visual-motor coordination; 2) Children *without* the nursery experience were more inhibited, somber, and rigid in their attitude toward the test situation; 3) All children learned on the problem-solving task under extrinsic reinforcement. However, only those children who had had prior nursery experience were able to learn under the condition of intrinsic reinforcement, in which a child has to learn from his own experience.

Milly Cowles, Rutgers, The State University, New Jersey. *Perceptual and Motor Development Training for Young Economically Deprived Children.*

This paper deals with:1) theoretical literature in visual perception and motor development in young children; 2) rationale for providing special training for kindergarten children in the two areas; 3) procedures for training which have been developed in the two areas; and 4) research design and methods which may be used to test the feasibility of a combined visual-perception and physical development training program.

E. Kuno Beller, Wilford Weber, and Edmund Amidon, Temple University. *Teacher Behavior and Intellectual Functioning in Deprived Kindergarten Children.*

The present study investigated relationships between. 1) gain in I.Q. score as a function of prior nursery experience, and 2) superiority in test performance of deprived children with such prior experience over children without such prior experience.

Three kindergarten classes of 98 pupils were observed. All children came from lower class, educationally deprived, mostly Negro families. The method of Interaction Analysis was used to measure teacher behavior. This allowed each teacher's interaction pattern to be classified as being direct (lecture, giving commands, criticizing) or indirect (accepting feelings, praising, accepting ideas, asking questions). Two verbal and two non-verbal intelligence tests were used to assess the impact of nursery experience.

Teachers were found to differ significantly in their behavior as measured by Interaction Analysis. Under the condition of indirect teaching, the impact of the nursery experience on the intellectual functioning of the children emerged most clearly. These results support the outcome of previous research on the positive effect of indirect teacher behavior and the impact of early educational experiences in deprived children.

SESSION 14

Paper session --STUDIES OF UNIVERSITY STUDENTS' ATTITUDES

Irvin J. Lehmann, Michigan State University. Changes in Attitudes and Values Associated with College Attendance.

This is a study of the impact of college on stereotypic beliefs, dogmatism, traditional values, and social, political and economic views.

A battery of cognitive and affective measures was administered to the freshman class entering Michigan State University in the fall, 1958. This same battery plus an elaborate experience inventory especially constructed for this study was administered at the end of four years. Ss were classified according to amount of college education and sex. Those Ss who withdrew from MSU were contacted and sent the test materials through the mail. The data were analyzed separately for males and females. Analysis of covariance and chi-square was used to test the hypotheses.

1) After making adjustments for initial status, for the males, there is a significant relationship between amount of college education and changes in traditional values. There is, however, no significant difference among the different (educational) groups in changes in stereotypic beliefs and dogmatism. 2). For the females, there is a significant difference among the groups in stereotypic beliefs and traditional values after four years. There is, however, no significant relationship between amount of college education and change in dogmatism. 3) There is a significant difference among the groups in certain social, political, and economic views, viz., admittance of Red China to the U.N., continuation of nuclear testing, medicare, etc. 4) Increased college attendance was accompanied by an increase in the proportion of Ss who became more a) tolerant of people differing in race and religion; b) respectful for the views and opinions of others, c) self-confident to deal with new problems, and d) inclined not to accept the Bible as a guide to modern living.

Edwin C. Lewis, Leroy Wolins and Julie Johnson Yelsma, Iowa State University. Patterns of Educational Attitudes among College Women.

Little empirical evidence is available concerning differences in educational needs and goals among college women. This paper reports the first in a series of studies designed to locate meaningful differences among college women in these areas which can in turn be related to predictive information so as to improve the relevance and value of their educational programs.

Ninety-five percent of a random sample of juniors, seniors, and first-year alumnae of the College of Home Economics, Iowa State University, responded to seven evaluative attitude questions concerning their reactions to seventeen courses in the Home Economics core program. Analysis of variance showed a highly significant person-course interaction, indicating that the subjects could be differentiated on the basis of their attitudes toward the various courses. Factor analysis of the attitude scores by the maximum likelihood method revealed six factors, which were defined in terms of the courses which loaded most heavily on each. It is believed that these factors are reflections of differences in educational needs and goals among the subjects and can thus be utilized as a basis for further research within this population.

SESSION 14

John A. Centra, Michigan State University. Major Field as a Variable in Student Perceptions of a Large, Complex University.

This study hypothesized that students within each college of a complex university have different perceptions of the total institutional setting; and that student perceptions of the total setting are related to student perceptions of their major field of study.

The instrument used to measure student perceptions of the total setting was C. Robert Pace's. Pace had identified five factors (or scales) which differentiated college and university educational environments: scholarship, practicality, awareness, propriety, and community. This study also used a subset of 80 items to assess major field perception on the same five factors.

A random sample of upperclass students was proportionally chosen (Winter term, 1965) from the ten colleges of a large university; 513 (80 percent) participated. Differences in total setting perception were analyzed by analysis of variance. For the two scales with the highest overall mean scores, i.e., the dominant features of the total environment, no differences were found between colleges. However, for two of the less dominant environmental features, differences between colleges were found, thus supporting one of the hypotheses.

Correlations between student perceptions of total setting and major field were significantly positive. Further study of the data revealed that each college group saw its environment as being more scholarly (higher on the scholarship scale) as well as less rebellious (lower on the propriety scale) than the total university.

Herbert J. Walberg, Educational Testing Service. Personality, Role Conflict, and Self-Conception During Student Teaching.

Research literature on beginning teachers suggests conflicts between the need to establish rapport with children and the professional role of teachers in maintaining authority. A theory that this personality-role conflict lowers self-conception was tested in a sample of 77 women. Ss took a semantic differential questionnaire before and after practice teaching, and rated themselves lower on 20 of a possible 20 items classified as favorable ($p < .001$) and higher on 8 of 8 unfavorable items ($p < .001$). Significant changes in individual items ($p < .05$) revealed the pattern of conflict more specifically. The phrases suggest less understanding of children (lower self-ratings on the items: "Always knows what mixes me up" and "Knows if we're trying"); lower expectation of pupil behavior and lower aspiration for self in the role of teacher (higher ratings on: "Makes the work too easy," "Doesn't have any rules," and "Let us be noisy," and lower ratings on "Makes us want to learn") and less rapport with the class (lower on "Makes school fun" and "Never makes us look foolish"). The changes in item means on the adjectives suggest other deleterious effects on self-conception. The changes imply self-depreciation on intellectual mastery and on the ability to present ideas effectively (lower ratings on the items "sure", "sharp", "plain", and "clear"); a sullied self-concept (higher on "sloppy" and "dirty"), and a dejected alienation (lower on "kind", "happy", and "familiar", and higher on "bouncy").

SESSION 14

James E. Greene, Sr., University of Georgia. Self-Reported Motives for Attending a Floating University.

This study reports analyses of certain self-reported motives for attending the University of the Seven Seas, First Semester, 1964-65.

The data available for analysis included two types of student responses: a) a one-page (or less) essay written during the first week of the semester responding to the question, "Why did you decide to attend the University of the Seven Seas?"; b) responses to a pre-categorized questionnaire requesting personal ratings of the "relative importance" of each of twenty specified motives for attending the U7S program (see Exhibit A). The essay responses were subjected to classification into several empirically-derived response categories based on the subjective judgment of the author.

The findings appear to justify among others, the following broad generalizations: a) both types of data indicate that the typical student was more strongly motivated by non-academic (e.g., world travel, self-directed tourism, socialization activities, etc.) than by strictly academic motives (e.g., improvement in academic skills, accumulation of credit, participation in student activities, etc.); b) marked individual differences exist in the reported motivations; c) inter-class differences in motivations are relatively slight.

SESSION 17

Paper session --NEW DIRECTIONS IN SCHOOL BOARD RESEARCH

Richard A. Rossmiller, University of Wisconsin. The Public Image of the School Board Role.

The report to be presented at this session will be derived from data gathered for USOE Cooperative Research Project 2371. The School Board As An Agency for Resolving Conflict. To obtain data concerning expectations for the school board held by the public, interviews were conducted with 1,794 randomly selected citizens in 12 Wisconsin school districts using a structured interview schedule. In addition to data concerning the school board role, the interview was designed to elicit expectations for the schools related to educational program, staff and pupil personnel, and business and plant management, as well as intermediate ratings of satisfaction and/or effectiveness for the board of education and the schools. Data concerning the background of the respondents were also obtained.

The findings to be reported include citizens' perceptions of: 1) Relative importance of the role; 2) Motivation for seeking the office; 3) Prerequisites for board membership; 4) Accomplishments and shortcomings of the board; 5) Operating procedures of the board; 6) Response to pressure groups; 7) Public involvement in policy making.

Findings will be reported for both the total sample and for the 12 school districts represented in the sample. Significant differences in expectations for the school board role among the 12 districts will be identified and significant differences in expectation held by various sub-publics will be highlighted.

Joseph M. Cronin, Harvard University. Changing the School Board in Great Cities.

An examination over seventy-five years of the ways in which the school board in the larger cities has evolved, stressing the events and social forces which have precipitated a variety of structural changes in city school boards. Preliminary findings on the implications of changing such variables as school board size and mode of board member selection will be reported.

Jay D. Scribner, University of California at Los Angeles. A Systems Analysis of School Board Action.

This study examines the applicability of a functional-systems framework to the performance of local school boards. A selected set of concepts was utilized in an attempt to produce a systematic description of (1) the different kinds of inputs entering the sub-system of the school board from its environment; (2) the kinds of political and governmental functions performed by the school board; and (3) the kinds of outputs emitted by the school board back to the environment.

The importance of this study is seen in the provision of a new approach to understanding the behavior of school boards. The next step should be to explore those concepts for which data was lacking and to expand a range of field studies to carefully chosen communities. Through these endeavors, the degree of reliability and validity of the conceptual classifications delimited in this study will determine which concepts will undergo change, be rejected or accepted. Those that do remain will contribute a set of conceptual tools to educational administration for studying the degree of specialization in administrative structures, and the frequency, kinds and styles of performance of functions by those structures in a variety of social settings.

Donald J. McCarty, Cornell University. *Myths and Reality in School Board Research.*

This research is concerned with testing a model which classifies the type of school leadership and the community power structure with relation to the system of public education.

The model may be summarized as follows:

Community Power Structure	School Board	Role of the Superintendent
Dominated	Dominated	Servant
Factional	Factional	Political Manipulator
Pluralistic	Status Congruent	Professional Advisor
Inert	Sanctioning	Decision-Maker

The general objective was to test the above model in terms of the frequency with which hypothesized relationships were observed. That is, are dominated community power structures usually associated with dominated school boards and with the servant role for the superintendent?

James W. Guthrie, Stanford University. *The Function of School Boards in the Mediation of Educational Demands.*

This study examines the processes by which local school boards mediate community educational demands and legislature mandated programs. Primary attention is given to the board of education as a structure for the maintenance of equilibrium. However, the study also includes information relevant to conditions of dis-equilibria, conflict, recall elections, etc.

SESSION 19

Paper session: LANGUAGE COMPOSITION

Anthony Tovatt, and Ebert L. Miller, Ball State University. Effectiveness of an Oral-Aural-Visual Stimuli Approach to Teaching Composition to Ninth Grade Students.

The paper will present the results of a research study with sixty (60) ninth grade students divided into an experimental class and a control class. The groups were stratified on the basis of sex, intelligence, and past English achievement.

Hypotheses tested focused on achievement and attitudes. First, there would be no significant differences between the experimental group (taught with an oral-aural-visual stimuli approach) and the control group (taught with conventional classroom procedures, but minus the special O-A-V approach) in achievement as measured by: 1) various standardized achievement tests administered at the beginning and end of the experiment; 2) judgments about samples of written composition taken at five different intervals during the experiment; and 3) a staff-constructed procedure for evaluating oral competence administered on three different occasions during the experiment. Secondly, there would be no significant differences between the groups on the basis of attitude change toward specific aspects of English as measured by two staff-constructed instruments.

Statistical techniques such as chi-square, "t" test, analysis of variance, and analysis of co-variance were utilized. In some instances the analyses justified rejection of the null hypothesis.

The paper will include a summary of the essential aspects of the Oral-Aural-Visual approach, which utilizes a special electronic device (a tape recorded adapted to provide immediate feedback to the user as well as pause and replay controls) together with materials and procedures designed to focus on multi-sensory involvement, immediate feedback, and immediate reinforcement.

Carl Personke, The University of Texas. The Effect of Systematic Instruction on Ability to Generalize in Spelling: A Comparative Study.

At the 1964 A. E. R. A. meeting, this writer reported on a comparison of spelling achievement in Scotland and the United States. It was demonstrated that Scottish children at all ages tested scored better than did the American children. Tests of phonetic rendition applied to misspellings indicated that the Scottish subjects tended to use phonetic generalizations more effectively. Frequency tables indicated that the Scottish advantage at age 14 was concentrated in the lower achievement ranges.

Since the Scottish advantage seemed to be the result of superior ability in applying phonetic generalizations, and since Scottish methods stress phonics and American methods do not, it was suggested that bright American children learned to make generalizations on their own; the duller children could not do this. To discover if this were the case, the samples were stratified into top, middle, and bottom 25% ranges. The same tests of phonetic rendition used in the original study were applied to misspellings of each stratified sample.

Comparison of most common misspellings yielded significant differences in favor of the Scottish children only in the bottom stratum. Tests of the readability of misspelled words, selected at random, yielded similar results. The difference of 7.8 words in the bottom stratum was the only significant difference discovered.

The results of this study indicate that (1) spelling generalizations can be learned, (2) direct instruction will be of most benefit to the poorer spellers, and (3) spelling achievement will improve with such instruction.

Geraldine E. LaRocque, Columbia University. *The Effectiveness of the Inductive and Deductive Methods of Teaching Figurative Language to Eighth Grade Students.*

The purpose of the study was to discover whether the inductive and deductive methods used by regular English teachers to teach figurative language would result in differences in eighth grade students' learning.

The 211 subjects were divided into high and low intelligence groups and randomly assigned to one of the two treatment groups or to the control. Following the pretested, highly detailed, structured lessons written for the experiment, two teachers taught both methods. The lessons were taped to determine method differentiation. Two days later the teachers administered tests of retention and transfer for which content validity was verified and corrected split-half reliability was calculated at .934 and .965 respectively. It was hypothesized that the mean scores of the students taught inductively would be significantly higher ($p < .05$) than the mean scores of the students taught deductively and that there would be no interactions involving sex, intelligence, method, or teacher. A four-way analysis of variance design was used to calculate the differences between the means of the experimental groups. The findings were: 1) Both experimental groups achieved higher mean scores than the control ($p < .001$). 2) The deductive method showed greater overall effectiveness ($p < .01$), but several interactions preclude a straight-forward interpretation of this main effect. 3) The most pervasive interaction was the one involving teacher, method, and intelligence. 4) The inductive method was consistently less effective than the deductive method with the low intelligence group.

William McColly, State University College, Oswego and Robert Remstad, University of Wisconsin at Milwaukee. *The Dimensions of Composition Annotation.*

It was hypothesized that the factor-structure of teachers' correction and criticism of compositions corresponds to that of composition itself as identified in several studies (Carroll, Diederich, Remondino), and it was assumed that knowledge of the dimensionality of annotation will lead to a better use of this instructional device.

A composition by each of a number of students in grades 9-13 was written on one of two topics (expository or narrative-descriptive) and under one of two conditions (impromptu or take-home). Typed verbatim and standard format, each composition was annotated by a classroom

English teacher. Each of 320 annotations ("one annotation" means the whole system of correction and criticism written on one composition), read independently of the composition on which it was written, was then judged by one English teacher from each of four judge types - college male, college female, school male, and school female - on each of thirty-two bipolar adjective scales within a four-day rating design counterbalanced as randomly as possible in terms of grade level, sex of writer, topic, writing condition, judge type, day, and day-session.

On the basis of the intercorrelation matrix generated by the mean ratings for each scale on each annotation, three different analyses were performed; image analysis, factor analysis, and incomplete components. These yielded three, four, and four general factors respectively. A degree of correspondence to the general dimensionality of composition appears, but certain important differences indicate that further factor-analytic studies of the phenomenon of composition annotation would be profitable.

Robert M. Rippey, University of Chicago. The Outcomes of Teaching English Composition by both Errorless and Dialectical Methods in Situations of High Teacher Involvement.

Students in two schools were taught highly structured and stylized maps for writing. These maps were taught using the errorless teacher and dialectical teacher cells of the Ginther Model of Instruction. Validity checks on the procedures employed by teachers were maintained using modified interaction analysis techniques. Differential effects due to the instruction in terms of eight (8) variables were inferred from student writing samples using trained readers who achieved a highly satisfactory degree of inter-rater reliability. This paper will present the differential effects of the two methods of teaching composition. Observations with respect to reactions of the students to the different methods of instruction will be discussed. Sound motion pictures of the classes in session will be included as a part of the presentation.

Since the experiment to be reported was rather complex, the phase of the experiment dealing with control of teaching procedures and the multivariate analysis of the differential outcome of instruction will be stressed.

SESSION 22

Paper session --MATHEMATICS AND SCIENCE LEARNING

Ralph H. Ojemann, E. James Maxey, and Bill C. F. Snider, University of Iowa. The Effect of a Programmed Sequence for Developing Probability Concepts at the Fifth Grade Level and its Contribution to a Theory of Instruction.

A programmed sequence of exercises based on an analysis of the learning product and nature of the learner and involving both verbal problems and simple experimental demonstrations was developed for teaching basic elements of probability thinking and their application to everyday physical and social phenomena. Equated experimental and control groups involving 54 fifth-grade pupils were utilized to test the effects. The program required seven sessions of 30 minutes each and involved 14 experimental demonstrations which the pupil could perform himself. A film program dealing with irrelevant material was prepared for the control group to take account of the "Hawthorne Effect".

Three pre-tests and three post-tests provided the data for equating groups and for measuring effects. Two of the post-tests involved making predictions when pertinent information was available. One post-test required predictions when input level was unknown but for which some data could be gathered during the course of the test. Significant differences between E and C groups were obtained on all three post-tests.

In addition to presenting the comparisons of E and C groups selected findings which appeared in the development of the learning program and which contribute to the development of a theory of instruction will be presented. Examples demonstrating how the analysis of the learning product and the nature of the learner were used in constructing the learning program as well as suggested studies involving the prepared program to test crucial issues in a theory of instruction will be discussed.

James R. Espenschied, and Philip R. Merrifield, Kent State University. A High School Physics Aptitude Study.

This pilot study was carried out to identify factors important in the Kent State University High School physics course. In the spring of 1965, scores on 22 variables were analyzed for 73 students enrolled in three physics classes. The students in all classes were somewhat above average on most ability measures and were all planning on college.

Eight of the variables were scores from tests representing specific cells in the structure-of-intellect model. This battery of tests was given to all classes early in March. Scores on the following 14 variables were available from school records: previous mathematics and science grades; verbal and non-verbal IQ; science and science-reading scores on the Iowa Test of Educational Development; five physics unit test scores; two laboratory grades; and the final examination grade. The unit tests in the study were those scheduled from January to the end of the school year and covered machines, kinetics, heat, light, and electricity. The laboratory grades were from written reports about the experiments on those units.

Frequency distributions and scatter diagrams showed that score distributions were nearly normal. From the Pearson-r intercorrelations of the 22 scores, seven principal components were extracted and rotated graphically to simple structure. Of the seven factors identified, four showed strong relationship to achievement in the physics course. These were given the names: convergent productive thinking (semantic); previous achievement; general reasoning; and figural cognition. The three factors on which the achievement variables did not have high loadings were: evaluation of semantic relations; productive thinking (symbolic and figural); and evaluation of symbols.

Ronald J. Raven, State University of New York at Buffalo. An Investigation into the Development of the Concept of Momentum in Primary School Children.

The purpose of this study was to determine the interrelations of the specific acquisitions of concepts that were postulated as being necessary for an understanding of momentum. One hundred and sixty subjects from the kindergarten, first, second, and third grades were asked to solve operationally these tasks.

With one exception, acquisitions of concepts were not related to sex differences or differences in mental age at each grade level. The success on transitivity of momentum at the first grade level was related to mental age. At least eighty percent of the first grade subjects were successful on both momentum tasks. All the subjects in the second and third grades were successful on the momentum tasks. The findings of this investigation agreed with Piaget's findings on the development of the conservation of matter, speed, and the proportion schema.

A comparison was made among the items with respect to their inference patterns, goal objects, and percepts. The use of different goal objects (speed and mass) did not cause dissimilarities in success on the momentum tasks when the percept and inference patterns were similar. However, the use of different goal objects (speed and mass) was associated with differences in success on the two proportion tasks when the percept and inference patterns were relatively similar. The proportional use of matter had more successes than the proportional use of speed.

Using the momentum tasks, an attempt was made to extinguish the subject's judgment based on transitive thought and his judgment based on immediate perception. More students extinguished their transitive reasoning.

Mary Budd Rowe, Columbia University. Some Properties of Context-Learning in Elementary Science.

The properties of context-learning differ in many respects from the properties of concept-learning. The paper presents data from portions of an empirical study done to determine characteristics of context-learning in elementary science, as it actually occurs - or fails to occur - in a sample of first grade boys. The paper stresses the distinction between the probability characteristics of the two complementary modes of learning and makes the point that context-learning is largely an ignored process in psychological research.

Probability characteristics of context-learning differ from those of concept-learning. Successful problem solvers employed patterns of redundancy that can be conveniently described as network sequencing of concepts and acts. Unsuccessful context-learners also employ a form of redundancy but the form may be characterized as repetitive. In contrast to protocols exhibiting network sequencing, repetitive redundancy protocols indicate that the student gains relatively less information per operation and error correlation from step to step is high. Some protocols of successful context-learners appear to be amenable to Markov treatment. Patterns of context-learning exhibited by subjects showed shifts in the direction of network sequencing as subjects accumulated experience in a variety of different but related contexts.

William F. White, University of Georgia. Teacher Motivation Cues and Anxiety in Relation to Achievement Levels in Secondary School Mathematics.

One-hundred-eighty-five high school junior and senior students were classified as underachievers, overachievers, and achievers, by percentile rank positions on CEEB scores and achievement tests in mathematics. An examination was made of the relationship of facilitating and debilitating anxiety and four teacher motivation cues to three levels of scholastic achievement. Girls were more sensitive than the boys to the motive-arousing cues provided by their teachers, yet differential sex and group reactions were reported in the perception of teacher cues. Mean scores for girls on the achievement variable are consistently lower across groups. Although there was no significant difference among underachieving, overachieving and achieving students in perceiving affiliative cues in their teacher, nevertheless the interaction of group and sex was significant at the .01 level. When the students were examined for the perception of structure, organization, and control in teacher motivation, girls seem to be more stimulated than boys by the controlled, conformed, behavior of the teacher. Female students were more sensitive to corrections by the teacher as well as announcements of tests and immediate feedback of the results. Facilitating as well as debilitating anxiety was a determining variable for both sexes on all three achieving groups.

SESSION 23

Paper session --MOTIVATION AND LEARNING

Jason Millman, Cornell University and Robert Stake, University of Illinois. Motivation and Rate of Learning: A Test of Carroll's A Model of School Learning.

In A Model of School Learning, John B. Carroll proposes that degree of learning depends upon a learner's aptitude (the amount of time needed to learn the task under optimal instructional conditions), his ability to understand instruction, perseverance (the amount of time he is willing to engage actively in learning), the time allowed for learning, and the quality of instruction. Motivation , desire to learn, is related to the factor of perseverance in that the greater the motivation, the greater the amount of time the learner is willing to spend actively in learning, all things being equal. That is, motivation is relevant chiefly as it affects the sheer amount of time that the learner will actively orient to the learning task.

It follows from the model that if a learner has sufficient perseverance to spend the amount of time that he needs (given the instructional setting) to learn the task, increasing his motivation or desire to learn will not alter the time needed to learn the task. Many observers would expect something else - that increased motivation would result in shorter learning times. The purpose of our study is to seek empirical resolution of this issue.

Several learner groups and paired-associate learning tasks will be used. Time to criterion will serve as the learning measure. A fast presentation rate together with the requirement of much overt response on the part of the learners will be employed to help insure that the learner is continually spending time on the act of learning. Motivation will be manipulated two ways - by changing the probabilities of reward and by changing amount of reward.

Robert J. Berger, University of New Mexico; Robert L. Baker and Richard E. Schutz, Arizona State University. The Effects of a Response Contingency on Responses to Audio and Visual Stimuli.

This study investigated Premack's proposition that the response characteristics of events can be used to establish reinforcement contingencies.

Thirty fourth-grade subjects were allowed to respond by pressing one of two switches. One switch resulted in an eight-second presentation of auditory material via a tape recorder, the other resulted in a similar presentation of visual material via a filmstrip projector. The frequency of each response was recorded during a twenty minute session. Subjects were randomly divided into "contingent response" (CR) and "free response" (FR) groups. The FR group was given another session exactly like the initial period, whereas CR subjects were given the response contingency. For subjects who initially responded with more auditory responses, the contingency consisted of requiring one visual response prior to each auditory response. The reverse contingency was given to subjects who initially made more visual responses. Thus responses of initially greater frequency were used to reinforce those of lower initial frequency. The effect of the contingency was determined by increases in the frequency of lower -rate responses.

The mean frequency of lower-rate responses for the FR group decreased by 7.4 from the first to the second session. For CR subjects, the mean frequency increased from 30.4 to 60.3, an average increase of 29.9 responses. The increase is statistically significant beyond the .01 level.

The results support Premack's proposition, and suggest that various instructional strategies be attempted in the classroom which set up a reinforcement contingency based on differential response rates.

Edward Levonian, University of California at Los Angeles. Retention of Information in Relation to Arousal during Continuously-Presented Material.

Utilizing stimulus material presented in trials separated by fixed time intervals, Recent studies at the University of Michigan reveal that material learned during high autonomic arousal tends to show reminiscence, whereas material learned during low arousal tends to show forgetting over time. It is the hypothesis of the current study that this same phenomenon will emerge for instructional material presented continuously.

On the day that a traffic safety film was scheduled for showing in four high school driver education classes, skin resistance measures were obtained from each of 83 pupils during class showing of the film. A 15-item questionnaire was administered immediately after the showing, as well as one week later. Each item pertained to either auditory or visual information presented at a specific point in the film. With this as the midpoint of a 1-minute interval, high arousal for a given pupil was defined as a resistance decrement larger than 0.5 SD of his 16080 resistance values standardized to unit variance.

Results: Utilizing independent entries, the 2 x 2 distribution of arousal (high or low) vs. type of retention (reminiscence or forgetting) was significant at the .001 level, and in the predicted direction.

Wayne Otto, University of Wisconsin. Anxiety and Motivation as Factors in Inhibitory Potential.

Studies of the effect of different levels of motivation upon the inhibitory potential of good and poor achievers have yielded ambiguous results. In an attempt at clarification, this experiment tested the hypothesis that subjects' anxiety level would affect their tendency to accumulate reactive inhibition.

The Children's Manifest Anxiety Scale was used to identify 120 pupils in Grades 4, 5, and 6 who scored high (top 20%) or low (bottom 20%) on the scale. Equal numbers of subjects were given instructions intended to evoke high or low motivation for the inverted digit printing task used to quantify reactive inhibition. Analyses of variance were used to compare the performance of high and low anxious subjects with high motivation and high and low anxious subjects with low motivation. The expectation was partially confirmed. With highly motivating instructions, high anxiety subjects accumulated more reactive inhibition than low anxiety subjects; but with low motivation the high and low anxiety groups did not differ. Apparently anxiety level becomes a significant factor only after a critical level of motivation has been reached. Nevertheless, the suggestion is that both anxiety and motivation must be considered in making predictions regarding inhibitory potential.

SESSION 23

Eugene J. Jabker, Miami University. The Motivational Effects of Non-Congruent Feedback.

The assumed functional relationship between certain patterns of teacher behavior and students' performance of required work (Cogan, 1956;1958) was experimentally investigated by the writer.

One hundred ninety-eight 5th and 6th grade male students were randomly assigned to one of nine individually administered experimental treatments which consisted of dyadic-combinations of evaluative feedback varying in degree of connotative congruency. The degree of congruency was effected by combining positive, negative and/or neutrally connotative evaluation from experimenter and non-experimenter sources. The appropriate treatment was repetitively administered to each student after the completion of three similar experimental tasks. Following the last experimental task, each student was requested to perform a "required" new task. The quality of the performance on the new task constituted the criterion data.

In general, the null hypotheses could not be rejected, casting some suspicion on Cogan's conclusions; however an unexpected finding occurred (as it had previously in a pilot study) which is the basis for the proposed paper. That is, students who received non-congruent feedback performed at a significantly higher level than students who received congruent feedback.

SESSION 24

Paper session --COUNSELING AND COUNSELOR TRAINING

Agnes G. Rezler, Northern Illinois University. *The Use of the Holland Vocational Preference Inventory (VPI) in conjunction with the Kuder Vocational Preference Record (VPR) in counseling high school girls.*

The first purpose of this study is to examine the nature and extent of the relationship between the Kuder VPR and the Holland VPI, a personality test composed of occupational titles. The second purpose is to examine the joint contributions of these two tests to the vocational counseling of high school junior and senior girls.

The devices were administered to 250 junior girls and 265 senior girls in a suburban parochial high school. Significant positive correlations were found between the following Holland and Kuder scales: 1. Realistic personality with outdoor and mechanical interest; 2. Intellectual personality with outdoor and scientific interests; 3. Social personality with social service interest; 4. Conventional personality with computational and clerical interests; 5. Artistic personality with artistic, literary and musical interests; 6. Self-controled personality with persuasive and clerical interests; 7. Aggressive personality with literary interest; 8. Masculine personality with mechanical, computational and scientific interests; 9. Status seeking personality with persuasive and literary interest. Significant negative correlations were also found between several personality and interest scales.

Conclusions: 1. While significant and meaningful relationships exist between the Kuder and Holland scales, the relationships are low enough to warrant the use of both instruments with the same group of people without duplicating information. 2. The joint contributions of the Kuder and Holland tests can be utilized to broaden the vocational horizons of high school girls and to help them to find vocations that will fit their personality and interest patterns.

L. M. Bennett and R. G. Peckens, Texas Woman's University. *Effectiveness of Communication of Secondary Level Guidance Counselors*

The purposes of this research were to determine to what degree 1) counselors effectively communicated test interpretation to counselees, and 2) counselees essentially understood an adequately stated test interpretation.

A selected group of 25 full-time secondary counselors were used for the first part of the research. In order to ascertain whether counselees understood an adequately stated test interpretation, an entire senior class was used as subjects for the second part of the research.

The following results were obtained: 1) Counselors spent undue percentage of session time in irrelevancies; 2) Counselors used terminology unfamiliar to the student and did not attempt to clarify technical terminology; 3) Counselors avoided discussing certain test results, especially those showing counselors in an unfavorable light; 4) Counselors tended to avoid discussing intelligence test scores whether favorable or not; 5) Counselors tended to produce sessions that were not organized nor planned. Their haphazard approach tended to reduce their effectiveness; 6) When the counselors were pressed for additional explanation, the continuity of the session was disturbed to the point that the counselor tended to repeat previously given material; 7) Test results showed that counselees did not understand an appreciable amount of the material given during the interpretation session; 8) The counselees lack of understanding did not prompt them to ask for clarification.

Ernest Spaights, University of Wisconsin, Milwaukee. Social Reinforcement in the Individual and Group Counseling of College Students.

The purposes of this study are to: 1) determine the effects of social reinforcement in short term individual and group counseling situations on accuracy of self-perceptions of college students; 2) ascertain the relative effectiveness of social reinforcement in individual and group counseling in effecting growth in the areas of critical thinking and attitudes toward teaching; 3) and compare the degree of counselee satisfaction with individual and group social reinforcement counseling. The subjects for this study were students enrolled in beginning teaching education course at the Ohio State University during the 1964-65 academic year. The total study involved 78 students, 26 in each of the two experimental groups (individual and group counseled) and 26 in the control group. Pre and post Minnesota Teacher Attitude Inventories (MTAI), Watson-Glaser Critical Thinking Appraisals (W-6CTA), and Student Self-Rating Sheets (SSRS) were administered. The fourth instrument, the Counseling Evaluation Sheet (CES) was administered to the experimental groups at the post-counseling stage. The investigator employed verbal conditioning as a counseling technique with the experimental groups while the control group received no counseling.

John D. Krumboltz, Stanford University, Lawrence E. Sheppard, Jefferson Elementary School District, Daly City, California. Job Stimulation: Effect on Career Exploration and Interests

The purpose of this study was to test ways of motivating high school students to explore their own career possibilities and the relevant educational and occupational information. It was hypothesized that tenth grade students who were given the opportunity to solve some simulated vocational problems would explore possible career opportunities and engage in various interest-indicative activities to a greater extent than equivalent subjects given two different types of control procedures.

A total of 500 high school students were randomly assigned to each of three treatment groups: 1) The Problem-solving group. This group received an "Accountant's Kit" which posed a problem like those faced by accountants which the majority of high school students could solve successfully; 2) The occupational-information group. This group was given some descriptive information concerning the occupation of accounting and a description of the problem that was presented in the problem-solving group but were not asked to solve the problem; 3) The control-information group. This control group was given only general information about occupations and the importance of planning one's future career.

Preliminary analyses of the results indicate the following: 1) Subjects in the problem-solving group reported a greater interest in the field of accounting than did the other treatment groups; 2) The problem-solving subjects requested information about a greater number of occupations than did the control-information subjects; 3) The control-information subjects requested information of a more general nature than did the subjects in the other two groups; 4) Subjects in the problem-solving and in the occupational-information groups requested more information about accounting than did subjects in the control group.

James H. Beaird, Oregon State System of Higher Education, Monmouth, Oregon. Audiosimulation in Counselor Training.

The purpose was to provide evidence relative to the effectiveness of a counselor training experience which used a simulated counseling interview as the instructional vehicle.

A simulated interview with a 7th-grade girl was developed, using all logical counselor responses to the girl's verbalization of her problems, and presented to experimental subjects by means of tape recordings. Each subject interacted with the prerecorded client verbalization, his response determining the simulated client's next verbalization. Twenty-eight graduate students majoring in guidance at Oregon State University were randomly assigned to two groups: 1) an experimental group receiving two hours of individual instruction with the simulated interview, and 2) a control group paired randomly and devoting two hours to role-playing conseling interviews, alternating between counselor and client roles. Prior to the training, all subjects participated in individual interview settings with a role-playing "client", and performance was scored, providing a pretraining measure of counseling performance. Counseling performance was evaluated: 1) by means of a ratio of acceptable responses (responses which tend to facilitate effective statements by the client or lead towards such statements) made by the subject to all responses made by him, and 2) by differences in the subject's use of acceptable responses in pretest and post-test interviews.

1) Subjects in the experimental group demonstrated a significant gain in their performance. 2) Subjects in the control group demonstrated a small but statistically nonsignificant decrease in their performance. 3) Subjects in the experimental group showed significantly more gain than did subjects in the control group. 4) When performance in the pretraining interview setting was controlled, there was a significant difference between experimental and control groups in post-training interview performance.

SESSION 26

Paper session -- RESEARCH ON THE TEACHING OF READING

Nila Banton Smith, Glassboro State College, New Jersey. Trends and Developments in Reading Research, 1884 to 1966.

This study involved historical research to ascertain changes in quantity, quality, topics treated, designs and statistical techniques used in conducting investigations in the field of reading in the United States from 1884 to the present time. Period: 1884-1910. Beginning of Reading Research. The first studies were laboratory studies, physiological or psychological in nature. Period 1910-1925. Initial Period of Emphasis upon Scientific Investigation in Reading. New instruments of measurement and new statistical techniques were developed. Most of the studies had to do with surveys and standardization of tests. Period: 1925-1935. Expansion and Intensive Application of Research. Topics were wider in variety and more sharply restricted, experimental techniques were improved, and interpretation became more discriminating. Period 1935-1950. Reading Research Has a Setback. During the war years, studies were reduced in number. New topics of high interest were: reading readiness, reading in content subjects and sociology of reading. Period of 1950-1965. Reading Research Grows in Quantity and Quality. Studies were greater in number and variety. Improvements were evident in significance of topics, in design, and in the use of several new statistical techniques.

Joel Strandberg, John D. McNeil, and Evan R. Keislar, University of California at Los Angeles. Experimentation in Teaching Young Children to Read Using a "Talking Book" System.

A unique "talking book" system has been devised at UCLA. This system, including a talking book and a nagmetic book reader, promises to meet special problems in the teaching of reading, such as the development of sight-sound correspondence of letters and words and particular problems of motivation. Some of the characteristics of the system (e.g., the simultaneous presentation of words and sound and the possibility of attaining mastery of the sound of any printed word during its initial encounter) make this system theoretically and practically important.

Evidence regarding the feasibility of instructing young children with the "talking book" system has been collected in experimental studies with kindergarteners. Results showed that children indeed were able to read combinations of words taught in the program as well as new words. The pupil response and excitement for this system were noteworthy.

Willavene Wolfe, and Bernice D. Ellinger, Ohio State University. The Teaching of Critical Reading: An Observational Study.

There were two specific purposes of the study: 1) to determine if some lessons for teaching critical reading were more effective than others in producing critical responses, and 2) to determine if certain categories of teacher behavior produce more critical responses from pupils than others.

The observational sample included 30 elementary classroom teachers, five at each of the first six grade levels, who had been identified as teachers who were now teaching critical reading. An observation scale was first developed for the specific purposes of this study.

A definition of critical reading was given to the observation teachers so that the project staff could be reasonably sure that they would observe lessons in this specific area of reading. However, the definition was kept fairly general to allow for expansion by the teacher. All teachers were observed twice: 1) teaching a lesson which they had designated as one providing for critical reading, and 2) teaching a critical reading lesson prepared by the project staff. An analysis of the data was made employing chi-square as the test of significance.

The data supported the position that: 1) certain types of lessons were more effective than others in teaching critical reading, and 2) specific categories of teaching behavior did produce more critical responses than others.

Robert E. Chasnoff, Newark State College. Initial Teaching Alphabet and the Traditional Alphabet in First-Grade Reading.

Seven pairs of matched classes of first graders were taught reading, writing, and other language activities by a single teacher. In each pair of classes, one class was taught in a classroom where only the Initial Teaching Alphabet (i. t. a.) was used. The matched class in each school was taught in a classroom where only the traditional orthography was used. The t-test statistic was used to test hypotheses with the following results: 1) The Stanford Achievement Test, Primary I Battery, Form W, transliterated into i. t. a. was administered to the experimental population. The same test in the traditional orthography was administered in the control group. A difference in favor of the experimental group was found at the . 01 level of significance. 2) All Ss were tested a month later on tests in the traditional orthography on Form X of the Stanford Achievement Test and on the California Reading Test, Lower Primary, Form W. The means of the experimental and control groups were not significantly different. 3. Four judges rated writing samples on a five point scale, using established criteria for rating effectiveness of communication. Samples were typed in traditional orthography on cards, with spelling corrected. Judges had no way of knowing which alphabet Ss used. Inter-judge correlations were made. The differences between the means of ratings was significant in favor of the experimental group at the . 01 level.

SESSION 31

Paper session--LINGUISTIC FACTORS IN VERBAL BEHAVIOR

Wilbur S. Ames, University of Missouri. A Study of the Process by which Readers Determine Word Meaning through the Use of Verbal Context.

The purpose of this study was to gain a better understanding of the process of using verbal context as an aid in determining word meaning. The introspective method of investigation was employed to study the thought processes of twenty graduate students who had been directed to use contextual aids to determine the meanings of simulated words. A total of 556 contextual situations and reader's oral interview responses were analyzed for the purpose of classifying the contextual aids that had been successfully used. The data were also analyzed to determine to what extent the use of contextual aids were related to the form classes of the unknown words and to identify predisposing conditions for the effective use of contextual aids. An analysis of the readers' responses to the 334 contextual situations in which they were judged to have attached adequate meanings to simulated words yielded a fourteen-category classification schema. An independent judge was used to determine the reliability of the investigator's judgements in both determining the correct responses and in categorizing the responses.

A relationship of statistical significance was found between correct responses to the simulated words and the form classes of the words. Both correct and incorrect responses of the subjects seemed to indicate that knowledge of word order and syntax, and background of experience served as important predisposing conditions to the effective use of context in deriving the meanings that might be attached to simulated words.

Daniel H. Muller, University of Wisconsin at Milwaukee. The Effect upon Pronunciation and Intonation of Early Exposure to the Written Word.

The study was designed to test the following hypotheses:
1) Exposure to the written word accompanying audio-lingual drill during the early stages of foreign language instruction has an adverse effect upon pronunciation. 2) Exposure to the written word during this stage of instruction has an adverse effect upon intonation.

One-hundred-forty-four students naive in the target language, Portuguese, were randomly assigned to two groups. The only significant variable was exposure of experimental group to printed forms of patterns being taught aurally and orally. Instruction carried out for six weeks prior to testing for differences in pronunciation and intonation skills.

Students' tests were taped and judged for pronunciation and intonation proficiency by a panel of raters. Using Kendall's coefficient of concondordance, the individual rater reliability was found to vary from 89 to 93 and the inter-rater reliability from .84 to .91.

The groups were found to differ significantly at the .01 level on poorly pronounced words and well pronounced words, and at the .05 level in intonation. All differences favored the control group.

Under the conditions of the study, both hypotheses can be accepted.

SESSION 31

Robert E. Davidson and Steve Lynch, University of California. *The Role of Syntactic Units in the Recall of Anagram Sentences.*

Edward Sapir has described the sentence as "topic" plus "commentary" -- a subject-predicate juxtaposition. It may be that the syntactic units which make up the subject or predicate are equally powerful in making the entire utterance comprehensible to a language user, but many writers on language emphasize the role played by the predication units. For example, Razran (1952) reported "preponderant effects of verbs and direct objects" when sentences are conditioned ,while "small weights" may be ascribed to the subject of the sentence. The role played by syntactic units and other variables in facilitating the recall and organization of anagram sentences was explored in two experiments.

In Experiment I, subjects (*S*s) heard 9 anagram strings of words, and their task was to recall and organize the strings into meaningful sentences. Independent groups of *S* s were *cued* by different syntactic units --article, subject, direct object, and verb. For two dependent measures, number of words recalled and number of sentences organized, the ordering of the means was identical -- article$<$ subject$<$ object$<$ verb. The ordered hypothesis was tested by means of the L statistic and was found to be significant at the .01 level.

Experiment II was a factorial design with repeated measures on two of four factors. The independent groups factors were ability (established by a pre-test) and syntactic units. The repeated factors were reversibility (subject and object syntactically and semantically reversible) and word frequency (Thorndike-Lorge). Analysis of the simple main effects with contrasts indicated that the syntactic units play different roles when interacting with ability groups and with other factors.

Joan L. Prentice, Indiana University. *Response Strength of Single Words as an Influence in Sentence Behavior.*

It was proposed that, in natural language, a communcation is begun with a verbal unit having high response strength (RS) at that particular time, and is concluded with some construction which will convey the intended message. If RS does, in fact, influence word order, and consequently sentence construction, in natural speech, it follows that *S*s would find it easier to learn sentences where word order corresponds to RS. Fifty-two *S*s were randomly assigned to one of four treatments. Twenty-six *S*s were assigned at random to one of two comparison groups. Active or passive sentences were learned in response to nouns. In one set of either construction, the first noun was elicited as primary word associate to a stimulus noun, and in re-paired sets the second noun was so elicited. Comparison groups learned the same sentences in response to unrelated stimulus nouns. Results indicated that sentences beginning with high RS nouns are easier to learn than sentences ending with high RS nouns. By inference, word order in natural speech is a function of RS, where verbal operants at high RS tend to be emitted early. Error analysis provided weak evidence for change in word order attributable to RS. The active construction appeared easier to learn than the passive construction, and syntactic errors tended to be in the direction of the simpler construction.

SESSION 32

Paper session --PERSONAL VARIABLES AND THE TEACHING PROCESS

Bernard R. Corman, Michigan State University. The stances of Teachers Toward Teaching

We define "stance" as the pattern of a teachers' beliefs about the purpose of her activity, her view of children, fellow teachers and administrators, and her commitment to teaching.

A taxonomy of stance-types based on the last of 10 depth interviews (median length = 2 1/2 hours) with each of 64 teachers over a three year span was derived. The exit protocols were rated independently by three judges using a 54 item set of scales. McQuitty's pattern analysis techniques were employed to identify seven stance-types under the decision rule that each teacher must be more like those with whom he is typed than any member of any other type. Descriptive labels assigned were: Child focusers, Pragmatists, Task focusers, Pragmatists, Task focusers, Contented conformists, Timeservers, Ambivalents, and Alienated.

Comparisons were then made to determine if the types discriminated independently of the material used in their derivation. No significant differences could be inferred on sex, age, marital status, or grade level taught. But significant and supporting differences were discerned on the Minnesota Teacher Attitude Inventory, on the Projective Draw-a- Teacher and on the Edwards Personal Preference Schedule. Finally, the responses of stance-type members were compared on four hypothetical situations involving decision-making in dealing with parental objections. The types did not differ in their actual decisions, but were markedly different in the extent to which they avoided a decision, appealed to the child's welfare, their own expertise or to abstract principles. Thus similar decisions carried quite different implications.

T. R. Husek, University of California at Los Angeles. Some Effects of Dogmatism in Elementary School Principals and Teachers.

Milton Rokeach's conceptualization of the dogmatic personality was used to generate the following predictions: 1) Principals, as authority figures, are exposed to complying behavior from dogmatic teachers; there is a positive relationship between principals' ratings and teachers' dogmatism scores. 2) Teachers are exposed to different behavior from dogmatic peers; there will be a negative relationship between peer ratings and teacher dogmatism. 3) Teacher-dogmatism will be related to descriptions of positive and negative authorities.

In 44 elementary schools 714 teachers and 44 principals completed the Dogmatism Scale, a teacher decision-making power scale, a sociometric ranking of co-workers, and a Semantic Differential with 15 scales and 15 potential school authorities as concepts. For 26 schools the investigator described the staff. The data were analyzed school by school, across groups of schools and also across all the schools. With a few minor exceptions, little support was found for the original predictions. However, the correlations (.67 to -.71) between teacher dogmatism scores and principal's ratings for the 44 schools indicated that the area is too complex for across-school analyses. When the 26 schools for which ratings were available on a "happy-troubled" scale, negative correlations between principal's rating and teacher-dogmatism were associated with "happy" schools, positive correlations with "troubled" schools. This relationship was significant beyond the .01 level.

Wayne K. Hoy, Oklahoma State University. Dogmatism and Pupil Control Ideology.

The school has been defined by R.O. Carlson as a service type organization similar to prisons and public mental hospitals in that on one hand clients have no choice in their participation in the organization; on the other hand, the organization has no control in the selection of clients. In such organizations client control is of particular importance.

In this study, pupil control ideology was conceptualized along a continuum ranging from "custodialism" at one extreme to "humanism" at the other. Custodialism was defined in terms of the traditional viewpoints of control while humanism connoted the conception of the organization as a community of human beings in which the varied needs of individuals are to be met.

Data for teachers (N = 805) were collected from a diverse sample of eleven school systems in Pennsylvania; the instruments were administered and completed at group faculty meetings. Data for principals (N = 168) were collected in the same fashion and, in addition, by mailed questionnaire in order to supplement the principals' sample.

In general, it was found that closed minded individuals as measured by the Dogmatism Scale were more custodial in their pupil control ideology than open minded individuals. In addition, organizational position was found to be significantly related to pupil control ideology. The pupil control ideology of principals was less custodial than that of teachers; further, secondary teachers were more custodial than elementary teachers in their pupil control ideology.

Leslie J. Wehling, Southern Illinois University at Edwardsville. An Exploration of the Organization of Teacher Conceptions of the Educative Process.

The problem of delineating system organization of teacher conceptions of the educative process was directed toward two questions: 1) What are the dimensions of conceptions that comprise system organization? 2) How are these dimensions ordered by individuals into conceptual types?

The data were collected by submitting a 118-item questionnaire in Likert format to a diverse population of 291 midwestern teachers and teachers to be. Six prominent factors emerged from a R-factor analysis of item intercorrelations, and these dimensions were confirmed: 1) an emphasis on content integration, 2) importance of subject-matter mastery, 3) teacher direction vs. pupil autonomy (a bipolar dimension), 4) consideration toward pupils (LBDQ type), 5) personal adjustment as an educational goal, and 6) a belief in emotional disengagement.

T-scores related to the conformed dimensions were submitted to Q-factor analysis. Clusters of individuals with factor loadings exceeding ± .50 were generated by three factors. Thus, three contrasting pairs, or six types of conceptual organization were identified. Significant Chi-square differences at the .05 and .01 levels were found between types and four face-sheet variables: present position, academic degree, age, and experience.

SESSION 34

Paper session:--CURRICULUM DEVELOPMENT

Elena M. Sliepcevich, National Education Association. *School Health Education: A Conceptual Approach to Curriculum Development.*

A conceptual framework for health education covering grades K through 12 was developed as a result of the efforts of a group of writers working together with teachers and supervisors from four tryout centers.

The framework includes a philosophical basis and a conceptual hierarchy of concepts viewed in three dimensions, physical, mental and social. A set of general student goals are categorized under Cognitive, Affective and Action Domains.

Using the framework as a basis, two sets of sample teaching-learning guides according to progression levels were developed. Included were extensive lists of additional readings, audiovisual materials and resource packets of selected authoritative materials to be used for the detailed content. The guides include behavioral outcomes, learning opportunities and evaluation procedures for the classroom. The tryout period was completed between January to June 1965.

Extensive feedback data were collected to evaluate the *procedures* of curriculum development, as well as the *product,* the curriculum materials. This paper will include the results of the evaluation of the procedures and only that part of the materials concerned with reactions to the materials from students, teachers, supervisors and field reviewers.

The experimental materials are being revised on the basis of the evaluation data and will serve as a prototype for the preparation of instructional materials for the remaining eight concepts.

Arthur R. King, Jr., University of Hawaii. *The Characteristics of Disciplines of Knowledge.*

The concepts of *disciplines of knowledge* and *structures of knowledge* have established themselves as vigorous debating points in the contemporary dialogue in curriculum theory. Bruner and other spokesmen for the disciplines and structural views of the curriculum promise astounding education advantages. Yet the terms remain highly metaphoric and vague. For example, the term "structure" has a rather precise technical meaning in mathematics, some currency in the physical sciences, but is untried and not yet useful in the social sciences and humanities.

If the potential of these concepts are to be exploited to the benefit of curriculum theory and practice, precising definitions and descriptive arrays need be developed. This study is one such attempt. The theme taken is the characteristics of disciplines of knowledge. Philosophical and architectonic views are set aside. The approach is analytic, using as data the statements of scientists, humanists, and others in the scholarly community about their fields of inquiry.

The paper reports the development of an analytic scheme of features of disciplines of knowledge. Their implications for further research and empirical testing, as well as their presumed implication for curriculum theory and practice, and educational evaluation will be developed.

SESSION 34

Thomas B. Goodkind, University of Connecticut. The Impact of State Curriculum Operations.

The study was designed to investigate the impact of the Illinois Curriculum Program upon local school district curriculum operations. The specific purpose was: to determine the extent to which State curriculum materials, in the form of subject area curriculum guides, were being utilized by teachers, administrators, curriculum directors, and county superintendents throughout the State; to investigate the relationship between utilization of the publications and such variables as type of school district, per pupil expenditure, teacher educational level and experience, availability of local curriculum materials, assistance from supervisory personnel, and perceptions of the purpose and usefulness of the publications.

Questionnaires were employed to survey the degree of usage of the publications by teachers and perceptions of their usage on the part of administrative and supervisory personnel. The sampling included ten per cent of the districts in the State representing 2/3 of the counties, and analysis of data was based upon replies from approximately 700 school personnel on all levels (43 per cent returns). Statistically significant relationships ($P < .05$) were found between teacher usage of the publications and a number of variables. Briefly, greater usage of State materials appeared to be related to the availability of help in their usage, type of school district, level of per pupil expenditure, and availability of local curriculum materials. Wide differences were found in the perceptions of teachers and administrators on the degree of usage and their importance.

Thomas L. Faix, Macalester College. Structural-Functional Analysis as a Conceptual System for Curriculum Theory and Research.

While recent curriculum revision projects have stimulated a new concern for learning in the separate subjects, it has become apparent that more attention is needed to the overall design of the curriculum in order to achieve a balance and better integration of knowledge. A major cause of this neglect rests with the lack of systematic knowledge about curriculum design. Variables have yet to be isolated and measured and laws remain undiscovered. This paper develops a conceptual scheme or framework of concepts that can be used as criteria or guidelines for more effective inquiry into curriculum phenomena. It presents a descriptive model or conceptual scheme that suggests the elements, structures, and functions of a curriculum system that are requisite for establishing, maintaining, and changing the system as a unified entity.

Much of what is taking place in the curriculum process may remain uncertain and uncontrollable, but by developing the field of curriculum inquiry more along scientific lines, this range can be reduced. Some rules of invariance will likely emerge. We are a long way, however, from theoretically derived and empirically confirmed propositions about curriculum phenomena. Indications are we will be at the present stage of general theoretical orientation for a number of years ahead. The conceptual system in this paper holds promise as a step toward more exact explanation and prediction of what happens in the curriculum process.

SESSION 35

Paper session -- RESEARCH METHODOLOGY

James M. Richards, Jr., American College Testing Program. An Analytic Model for College Effects

Recently increasing attention has been paid to investigating whether different colleges affect students differentially. The best "input-output" design developed by Astin is limited in that real college effects can be identified only by significant partial correlations, in which differences in "output" are controlled for differences in "input" or by discrepancy scores between actual and predicted output. Such measures are unreliable and subject to artifacts of various sorts. In order to improve on this design, a set of unifying concepts is needed for distilling the essential facts and portraying them quantitatively in a form that researchers and policy makers can use effectively. The proposed model is a step in meeting this need.

This model, an analog of the college effects process, is made up of two kinds of elements; boxes, representing pools of students distributed according to some (dichotomized) variable of interest, and lines, representing flows of students into and out of the pools. By quantifying each element in the model in terms of a probability, it is possible to write an equation stating "output" as a function of initial probabilities (input) and of probability of change. A consequence of the model, therefore, is that differential college effects are measured by different probabilities of change. A further consequence is that it is only necessary to control those student characteristics correlated with change, not all those correlated with input. The model is illustrated by application to artificial data.

Martin K. Chen, U.S. Public Health Service. The Matching Technique in Experimentation: An Evaluation.

There is a need to subject matching pairing techniques to a more rigorous scrutiny than has been hitherto undertaken in any single article on research design and statistical methodology. This paper attempts to show that many of the objections to matching commonly raised in books on experimental design and statistics can be overcome to a certain extent, but that there are other fallacies in statistical reasoning which are not generally realized and which militate against the indiscriminate use of this technique in experimentation. It is pointed out that when a priori knowledge about the relationships of the matching variable(s) and the dependent variable is lacking, the value of matching to the experimenter becomes dubious. It is demonstrated that when experimental subjects are matched on a variable not related to the dependent variable, not only is there no increase in precision but the error term used in the statistical test becomes suspect. Further, the pitfalls of manipulating a given supply of experimental subjects on a purely judgmental basis are pointed out to the unwary researcher in terms of the well-known properties of the statistical technique known as the analysis of variance. Proof is given, on an intuitive basis, that artificially equating the experimental groups with respect to mean and variance, while it is capable of reducing bias in statistical estimation, will also artificially inflate the error term with which the estimate is evaluated.

SESSION 35

Kenneth H. Wodtke, Pennsylvania State University. On the Assessment of Retention Effects in Educational Experiments.

Several problems in studying retention effects in research designs employing repeated posttest measures are discussed.

a) Differences among treatment groups in the amount learned have frequently been confounded with tests of retention effects. The confounding results from a failure to take into account the fact that the measure of the amount learned and the amount retained are usually positively correlated. If the investigator finds differences in the amount learned attributable to the experimental treatment, similar differences in the amount retained on a delayed test are expected as a result of the correlation. Failure to hold constant the amount learned,either statistically or by direct manipulation,has led to misinterpretation of retention test data.

b) The possible effects of regression towards the mean on retention data are also discussed. Investigators have frequently failed to recognize the possibility of confounding regression effects with the effects of forgetting.

The application of a factorial repeated measures design to the study of retention over time is described. In this design, the problem is one of comparing the slopes of forgetting curves. The particular application of the design provides a control for the effects of regression towards the mean. Other procedures were also discussed.

Examples are given of instructional treatments which may not be effective in increasing the amount learned put unit of time but which may have a most important effect on the rate of forgetting.

William A. Mehrens, Michigan State University. The Use of Probability Information in Educators' Decision Making.

The more widely accepted decision theory models suggest that people do not make decisions on the basis of the objective probabilities and monetary values of the possible outcomes. Rather, decisions are made on the basis of the subjective probabilities and utility values of the possible outcomes. Since decision making does appear to be related to the parameters of subjective probability and utility, it seems that a more microscopic study of these variables is in order.

This study used the SEU decision theory model in investigating the relationships between the independent variables of objective probability and a student's expressed interest in attending college, and the dependent variables of utility and subjective probability. This was done for both counselors and teachers.

It was found that: 1) Both dependent variables were affected by the independent variables. 2) The counselors' utilities were more affected than the teachers' utilities by the students' expressed interests in attending college. 3) This same differential effect did not hold true for the educators' subjective probabilities. In fact, counselors' subjective probabilities were less influenced than teachers' by the students' expressed interests in attending college.

SESSION 38

Paper session --, MODIFYING TEACHER BEHAVIOR

Richard E. Hodges, University of Chicago. A Theoretical Discussion of an Instructional Flexibility Training Model for Teacher Education.

One of the primary goals of elementary teacher education programs is to enable teachers to display controlled flexibility of teaching behavior appropriate to instructional settings or to the purposes of instructional situations. It is posited that the teacher who can purposely exhibit a wide range of teaching behaviors is potentially able to accomplish more effective teaching than does the teacher whose teaching behavior is limited. Recent studies of teachers' abilities to manifest controlled instructional flexibility, however, indicate that many teachers maintain rather rigid styles of teaching.

The rationale and structure of an elementary teacher education group is described in which explicit attempts are made to train prospective teachers to display controlled variability of teaching performance. This program has four stages of training for instructional flexibility in which students are helped to: 1) discriminate and analyze teaching behaviors using several frames of reference; 2) analyze their own teaching behaviors using these same frames of reference; 3) select and use teaching behaviors that normally are not a part of their instructional repertoire; and 4) measure their progress toward increased instructional flexibility.

The conceptual framework of this model of a program in instructional flexibility training is further elaborated through a description of the instruments used to assess students' progress toward controlled teaching behaviors.

Bruce R. Joyce, Manhattanville College of the Sacred Heart. An Evaluation of an Instructional Flexibility Training Model through a Developmental Study of the Teaching Styles of Elementary Teacher Education Students.

The paper presents the results of a developmental study of the teaching styles of elementary teacher education students. Tape recordings and films were made of classroom behavior over a one-year period. The verbal behaviors of the teacher education students were analyzed from several frames of reference to determine: 1) Whether flexibility of styles changed during the training period; 2) The developmental pattern in rewards and punishments, information - handling, managerial behavior; 3) The relation of personality variables from the (Conceptual Systems point of view) to changes in teaching style, especially changes in flexibility of style.

Significant changes in flexibility occurred, for both groups, with personality factors surprisingly effective predictors of change patterns. The analysis of developmental teaching patterns shows marked change in rewarding-punishing behavior, information-handling, and managerial behavior. Intellective and flexibility factors appeared to emerge.

The effectiveness of the instructional flexibility training model seems confirmed by the results, and the developmental has implications for teacher education procedures outside the model structure as well.

John B. Hough, and Richard Ober, Ohio State University. The Effect of Training in Interaction Analysis on the Verbal Teaching Behavior of pre-service Teachers.

The purpose of this research was to study the effect of instruction in interaction analysis on the verbal teaching behavior of pre-service teachers. Four hundred and twenty students were randomly assigned to twenty class-sections of a general methods course. Five treatment groups were used, and each group was composed of four class-sections.

Within treatments, all factors of instruction were held constant except the two experimental variables. The experimental variables were: 1) the means by which teacher verbal behavior was studied and analyzed; i.e., a) using the formal system of interaction analysis, or b) without the aid of a formal category system; 2) the means by which human relations concepts were studied, i.e., a) by means of the H. D. I. Relationship Improvement Program, or b) class readings and discussions of human relations concepts, or c) dyadic discussions of educational case studies.

During the 8th and 9th weeks of each quarter, all subjects planned, taught and evaluated a half-hour lesson. Trained, reliable observers categorized instructional behavior in all lessons using interaction analysis. Subjects taught interaction analysis were found to use more praise and encouragement, more acceptance and clarification of student ideas and more questions. Subjects who were not taught interaction analysis used more directions and criticism. In addition, it was found that subjects who were taught interaction analysis obtained more student-initiated responses while subjects not taught interaction analysis obtained more teacher-initiated student talk.

W. James Popham, University of California, Los Angeles. Relationships between Highly Specific Instructional Video Tapes and Certain Behaviors of Pre-service Teachers.

The research was designed to test the efficacy of four video taped instructional sequences in bringing about certain test behavior changes in prospective teachers. A series of 4 instructional video tapes were prepared dealing with the following topics: 1) Appropriate Practice; 2) Reinforcement Principles; 3) Discipline; 4) Learning Sets. These video taped programs were followed by a video tape post-test which assessed one's ability to perform the skill.

Group 1 will be the no-instruction group. Group 2 will view two instructional programs and group 3 will view the other two. Groups 2 and 3 will also receive additional instruction regarding the principles treated. Groups will be compared on paper-and-pencil tests, a video-tape post-test, and attitudinal measures.

SESSION 41

Paper session --CONCEPT LEARNING

John T. Guthrie and Richard C. Anderson, University of Illinois. *Stimulus Sequence and Concept Learning*

Previous studies have suggested that concept learning is facilitated when the exemplars of a concept are grouped together on successive trials. There is reason to doubt this conclusion since the method of composing the series in these studies probably confounded grouping with the number of irrelevant stimulus dimensions changing from trial to trial, another variable that may affect concept learning.

Two experiments were completed that compared constant series, in which stimuli containing the same relevant cues are grouped together on successive trials (e.g., red, red, red, red), with alternating series, in which one set of relevant cues appears on even trials while another set appears on odd trials (e.g. red, green, red, green).

Both experiments revealed a significant ($< .01$) interaction between type of series and the number of irrelevant stimulus dimensions changing from trial to trial. As predicted, learning was facilitated in constant series when the number of irrelevant dimensions changing was maximized. Within alternating series, the rate of learning was highest when the number of irrelevant dimensions changing was minimized. These experiments failed to show any overall advantage for constant series, in fact, the evidence tended to favor alternating series.

M. C. Wittrock, University of California at Los Angeles. *Replacement and Non-Replacement Strategies in Concept Identification.*

From mediated generalization theory, it was hypothesized that in a concept identification study with learning and transfer, the dependent variables treatment group means would rank from high to low: Non-Replacement, Replacement, and Control. In the Non-Replacement treatment, the children turned face down any card representing a concept just determined to be incorrect. In the Replacement treatment, the card remained face up. The Control group had no cards.

119 primary school children from two Los Angeles public schools were individually assigned at random to the three treatments. All groups received the same familiarization training, problems during experimental training, and examinations after training. Each training problem involved matching a top picture to one of two bottom pictures on the basis of color, size, shape, or number. Each child sat in a booth and viewed the problems projected on a screen. Through earphones, he heard E give instructions. The four cards lay on his desk before him. By pushing one of two buttons on his response panel, he received immediate feedback. All stimuli were automatically presented and responses recorded by automated equipment.

Four post-tests were given, learning with and without cards, transfer to new sizes, shapes, etc., and transfer to different concepts. On each of the four tests, the F ratio from the analysis of variance for a simple randomized design was statistically significant at the .01 level. The rank order of the means was as predicted with nearly all predicted differences among the means statistically significant. Apparently, young children can be taught to transfer a systematic way to eliminate incorrect choices from among four possible ones.

James G. Ramsay, University of Wisconsin. The Attainment of Concepts from Figural and Verbal Instances by Individuals and Pairs.

The purposes of the study were (1) to compare the performance of pairs to that of individuals on a concept attainment task, and (2) to determine the effect of two types of material, figural and verbal, on concept attainment. Subjects were 96 college students, 32 pairs and 32 individuals. Experimental materials consisted of two, multi-dimensional stimulus arrays of 64 cards with each card bearing one of two values of a color (red or green), number (one or two), size (large or small), shape (circle or ellipse), texture (spotted or plain), and number of borders (one or two). The information on one of the arrays was represented figurally; on the other array, it was printed in words. The task for the S or pair was, given an initial positive instance or focus card, to attain the concept by selecting other instances from the array which E would then identify as positive or negative instances.

Two dependent variables were used to assess performance: time to criterion and total number of card choices. For each dependent variable, a three-way analysis of variance was computed with size of group, type of material, and sequence of concepts as factors. From the analyses, it was determined that pairs made significantly fewer card choices than individuals and that Ss attaining concepts on the figural material took significantly less time than those on the verbal material. From the findings, it is suggested that pairs were more efficient in information processing than individuals and that information was more readily available on the figural material than on the verbal material.

Fred W. Ohnmacht, University of Maine. The Effects of Field Independence and Dogmatism on Reversal and Non-reversal Shifts in Concept Formation

44 Ss were administered the Dogmatism Scale (D) and a form of the Embedded Figures Test (EFT) which was used to measure Field Independence. Median splits on these measures served to identify groups of ten each, representing the four possible combinations of high and low when the two distributions are considered concurrently. Four Ss were eliminated using a table of random numbers to produce equal sized sub-groups. Half (5) of each sub-group was randomly assigned to the reversal (R) shift group, and the other half was placed in the nonreversal (NR) group. A card sorting task after the work of the Kendler's was constructed to serve as the experimental material. Ss were trained to sort the cards to a criterion state for an initial concept (15 consecutive correct sorts). Upon reaching the criterion for the first concept Ss were not told, but continued to sort with E changing

the feedback to conform to R or NR shift conditions according to a prior random selection. The design consisted of a 2 x 2 x 2 factorial with five replications in each cell. It was hypothesized that significant differences favoring the reversal shift condition, low dogmatic Ss, and field independent Ss would obtain. These hypotheses were confirmed at the .05 level of significance. It was further hypothesized that high dogmatic (HD) - field dependent Ss would have particular difficulty with both the R and NR shift conditions. The D x EFT interaction ($F = 2.92$) has an associated probability of .10. An inspection of the form of the D x EFT interaction suggested that the hypothesis has some merit.

Joe L. Byers and Robert E. Davidson, University of California. *The Role of Hypothesis Behavior in the Facilitation of Concept Attainment.*

Forty-eight Ss in each of two groups worked three successive conjunctive concept attainment (CA) problems. The Ss in one group were required after every trial to offer overt verbal hypotheses about the rule the E had used to categorize the stimuli into a positive instance and a negative instance subset. Ss in the other group offered hypotheses about the concept rule only when they believed they knew the rule. These instructions resulted in the first group, the hypothesis required (HR) group, offering hypotheses on all trials and the hypothesis not required (HNR) group offering hypotheses on only 25.7% of the trials.

Predictions about the differences in performance were based on a verbal mediation theory of problem-solving. It was predicted that Ss in the HR group were systematically required to practice a verbal chain which facilitated CA performance. This chain was described as a series of questions and answers relating to the status of the various stimulus attributes. The early acquisition and continued use of this chain should improve CA performance by reminding the Ss which attributes were relevant, which were irrelevant and which remained to be disambiguated. A subsidiary prediction from this theory was that the general inter-concept practice effects commonly noted in CA tasks should be reduced or absent in the HR group. This prediction implies that at least part of the learning-to-learn effect in CA is the result of the slow acquisition over concepts of verbal chains which serve to improve performance. Both predictions were supported by the data.

SESSION 46

Paper session --INSTRUCTIONAL VARIABLES AND CURRICULUM

John M. Long and Howard J. Barnhard, University of Arkansas Medical Center. Dictionary Buildup and Stability of Word Frequency Rank in Specialized Areas

This paper reports our study of word usage in a specialized area. We have analyzed about 100,000 words of running text of radiological patient data. Such studies are useful in information storage and retrieval problems and have overtones of implications for the educational process.

For example, the 25th ranked word was within 10% of its final position when about 8,000 words had been covered; the 50th word took 20,000 words; the 100th word took about 40,000 words; and the 500th word took 55,000 words to attain a 10% stability in the sample studied. Words ranked higher than 1000 were still quite erratic in their ranking after 100,000 words of text had been tabulated. It is not known if this relationship can be adequately represented by a linear equation or not.

It was also found that the top 25 words in rank by frequency of use accounted for about 46% of the total word usage. (The top ranked word "THE" accounted for about 10% of the words). The first 50 words accounted for about 56%; the first 75 words for 62% and the first 100 words about 66%. An attempt has been made to find a mathematical equation to describe the relationship between the frequency of use(or rank) and the information carried by the word. A simple inverse relationship does not appear to be adequate.

Hulda Grobman, University of Colorado. Achievement of Lower Ability Senior High School Students in BSCS Biology Special Materials, and Significant Variables Influencing Achievement.

The BSCS has prepared experimental materials for a one-year high school biology course for students whose reading and reasoning ability prevents their using regular 10th grade BSCS Biology materials. In 1964-65, SM materials were tested by 300 teachers with 15,000 students. All teachers attended a two-day training session. Classes include only those where the mean DAT (VR+NA) raw score is not above the 50th percentile. All students took beginning-of-the-year DAT (VR, NA, AR) and Processes of Science Test, end-of-unit tests, and an end-of-year comprehensive final and the Processes of Science Test. Several subgroups were selected each including half advantaged and half disadvantaged schools. Each subgroup was given one of the following beginning and end-of-year tests: DAT (VR, NA); Inventory of Adjustment and Values; a semantic differential Instrument; Metropolitan Reading Test; Cornell Test of Critical Thinking.

The report will include an analysis of student scores on subject matter and process skills and on other instruments, and an analysis of variables in the teaching situation, including whether school is advantaged or disadvantaged, background of teacher, ability composition of classroom. Procedure for analysis will be comparable to that in the 1963-64 Special Materials evaluation.

SESSION 46

Vernon S. Gerlach and Howard J. Sullivan, Arizona State University. Efficient Development of Pitch Discrimination.

In many curriculum areas simple discrimination tasks are desired terminal behaviors. In addition, many more complex objectives are mediated by discrimination learning. Laboratory studies have demonstrated the effectiveness of certain procedures for acquisition of discrimination behavior in animals and humans. If such training methods were generalizable to conditions in elementary and secondary schools, teachers would have at their command a relatively powerful means of implementing many of their objectives.

This study was conducted in a context related to the school music curriculum. The specific objective was the development in the S of pitch discrimination. It was hypothesized that errors are not a necessary condition for retention of an operant discrimination in humans. Experimental Ss were trained to select a correct symbolic (notational) representation of an audible stimulus (tone). Beginning with an octave interval, pairs of notes of successively decreasing intervals were presented to S, who pressed a button to indicate which one of a pair of notes represented the tone he had just heard. The performance of experimental Ss was compared with that of Ss who received equal practice with non-sequenced stimuli. The results extend the principles of errorless training reported by Terrace (1963) and Hively (1962) to an academic task.

Edward B. Fry, Lee Mountain, and Helma Heard, Rutgers-The State University of New Jersey. First Grade Reading Instruction Using a Diacritical Marking System, The Initial Teaching Alphabet, and a Basic Reading Series.

Many schools in the United States and Britain are starting to teach beginning reading using the ITA. The problem is: "Will children learn to read better in ITA classes rather than regular classes?" Partly to control for the Hawthorne effect two new methods were used. The DMS also regularizes the orthography, and looks novel. Is it any better than ITA or TO? To find out, 21 first grades were randomly assigned to the 3 methods, 7 classrooms per method. There were no significant differences between methods groups' initial IQ or Reading Readiness scores.

After 140 days of instruction (May), there were no significant differences on mean scores of any part of the Stanford Achievement Test. These sub-tests included Paragraph Meaning, Word Reading, Word Study Skills and Vocabulary. The spelling subtest showed ITA inferior if TO standards were upheld but no difference if ITA spelling was permitted. Further analysis is being done on sub-groups determined by sex, IQ, and socio-economic status and various individual tests.

A correlation matrix was made with 17 variables. One of the highest correlations, .81 was between IQ Raw Score (MA) and Paragraph Meaning. Thus MA predicts this type of reading success better than reading readiness test total or sub-test scores.

SESSION 47

Paper session --COMPUTERS IN MEASUREMENT AND RESEARCH

John F. Vinsonhaler, James E. Molineaux, and Bruce Rogers, Michigan State University. An Experimental Comparison of Computer Aided and Traditional Test Administration.

The purpose of the investigation was to compare the traditional versus automated administration and scoring of the SCAT, College Ability Test (verbal). The method involved the administration of the test to two carefully matched groups of high school and college subjects. In the control group, the test was administered and scored by an experimentor using test booklets and answer sheets. In the experimental group, the test was administered and scored by an IBM 7010 computer using an electric typewriter for stimulus presentation and response recording. The published, standardized directions for test administration were followed exactly in the control group. Standardized procedures were also followed in the experimental group except for an increase in control over the sequence of item presentation and response recording.

The major results of the experiment were that the automated form of the test yielded somewhat better test performance and lower test variability, but required more time for administration than the traditional test booklet form of the same test. Thus, the mean total test score was slightly larger, the variance significantly smaller, and the mean time on the test significantly larger in the experimental group than in the control group. The differences in performance and variability were interpreted to be the result of a tendency of the computer presentation to hold subjects' attention and to eliminate variability due to lack of control over the testing situation. The difference in time on test was attributed, in large measure, to the very slow rate of the typewriter mode of item presentation.

M. Clemens Johnson, University of Michigan. A Computer Simulated Classroom.

This paper discusses the design of a computer program written to simulate student learning in a hypothetical fourth grade class. A wide range of initial conditions can be simulated with the probability based model. Input data include two class characteristics, and four teacher characteristics, each of which is set at one of five levels. The class characteristics are: average scholastic aptitude; and homogenity of students. Teacher characteristics are: subject matter covered in class; explanational skills; motivational skills; and years of experience. An optional weight can be assigned to each teacher characteristic in the estimation of a composite teacher effect. In a particular computer simulation, computer generated data are printed for each of 20 "students". Output data for the individuals include chronological age; intelligence quotient; grade level for beginning of year, middle of year, and end of year; change score; and percentage change. Output data for the class include a high score; low score; mean; and standard deviation for each variable. In addition, correlation coefficients are computed between initial and final grade level, and between intelligence quotient and final grade level. The variability of output data; measures of central tendency; magnitude of change scores; and arrangement of data (together with student names) resemble in appearance a page from a teacher's record book. The exploratory program is written in the MAD language (Michigan Algorithm Decoder) and is processed on the IBM 7090 data processing system. Different sets of initial conditions may be simulated in a single computer run.

Harold F. Rahmlow, Washington State University. Computer Usage in Educational Statistics.

Because graduate students in education are typical users, not creators of statistical procedures, it is desirable to introduce the students to sound statistical procedures and effective data reduction techniques as rapidly as possible. This study attempted to provide an introduction to computer techniques as an integrated part of a course in educational statistics. The pure statistics portion of the course was data processing centered around statistical procedures previously studied.

Library programs were the basis of instruction with no attempt being made to teacher programing. A program (BIMD 16) of the Biomedical Computer Programs series was used as a working model. The students successfully completed a number of exercises in data reduction using the BIMD 16 program in conjunction with unit record equipment and an IBM 709 facility.

The cognitive aspects of the study were evaluated primarily through a use test in which the student was given a novel problem to solve which was solved satisfactorily. Affective aspects were evaluated primarily by the semantic differential technique. A favorable change of attitude toward the concept "statistics" was found (p = .05).

John J. Schurdak, IBM Watson Research Center. An Approach to the Use of Computers in the Instructional Process and an Evaluation.

An approach to the use of computers in instruction was developed and evaluated. In a computer course, students communicated with a 1440-1448 computing system through 1050 terminals, which have keyboards like typewriters. Three experimental groups learned a portion of a Fortran course by three treatments - by computer, programmed text, and conventional text. An achievement test was administered the day following completion of the course. The computer group's mean score was 11%, over one standard deviation, higher than the programmed text group, which in turn performed 6% higher than the conventional text group. Analyses of covariance, using as covariates scores from the Henmon-Nelson Tests of Mental Ability, disclosed that the means for treatments were significantly different and that the mean for the students in the computer group was also significantly higher than that for the programmed text group. There was least variation in achievement test scores among the students in the computer group, and the most variation in the conventional text group. The students' attitude toward the computer course and equipment, and toward the programmed text was good.

SESSION 48

Paper session -- STUDIES OF STUDENT CHARACTERISTICS II

Richard A. Rehberg, University of Oregon. Selected Determinants of Adolescent Educational Expectations.

Adolescent educational expectations are conceptualized as consisting of specific goal levels ranging from the non-completion of high school through graduate school. The specific goal an adolescent will select is considered to be a function of at least three variables: 1) the affective valence toward three educational functions: a) instrumental, b) expressive, c) symbolic; 2) the moral evaluation of education, i.e., as a social "good" or "evil"; 3) the facilities a) intellectual, b) economic, c) role performance, available to attain any given educational goal.

Sources of variation in these three variables were hypothesized to be: 1) social status, 2) achievement and independence training which are functions of family power structure which in turn is a function of family size, 3) education of each parent and parental educational disparities, 4) the influence of the adolescent's peer group.

Data for the study were collected in 1963 from a survey of 6000 students enrolled as sophomores in all public and parochial secondary schools in six Pennsylvania cities. This paper discusses analyses for the male data only.

Statistical testing of the various hypotheses permits the generalization that the proportion of adolescents actually expecting to enroll in four year colleges: 1) Varies positively with social status; and with each status category; 2) Varies inversely with family size; 3) Varies positively with intensity of parental achievement pressure; 4) Varies positively with the education of each parent; 5) Is greater than the stratum mean under the condition of maternal (but not paternal) educational superiority; 6) Varies positively with the proportion of the peer group which is college oriented.

Lee M. Joiner, Michigan State University, Edsel L. Erickson, Western Michigan University, Wilbur B. Brookover, Corwin A. Krugh and Natalie Sproull of Michigan State University. Perceived Expectations and the Educational Plans of Male High School Students: A Longitudinal Study.

The present investigation tested two major hypotheses: 1) Perceived educational expectations of parents, teachers, and friends are factors affecting student plans; and 2) Perceived educational expectations of parents are more highly related to student plans than perceived educational expectations of friends. The population in this investigation was an entire class of male Caucasian students followed through eighth, ninth, and tenth grades in a midwestern city. Magnitude of hypothesized relationships between the educational plans of students and the perceived educational expectations of others was determined through rank order correlation analysis while association of changes was analyzed by means of chi-square and contingency co-efficients. Tentative findings support the major hypotheses while additional analysis is concerned with the following questions: 1) What trends, if any, appear to be developing in the association of perceived expectations and educational plans from grades eight to ten? 2) Do increasing expectations exert as strong an influence on educational plans as do decreasing expectations? 3) How stable are the perceived expectations of others and the educational plans of students? 4) Do the perceived expectations of friends, parents, and teachers function independently of each other?

Carson M. Bennett and Daryl L. Dell, Ball State University and Maurice J. Eash, Hunter College. An Exploratory Study of Perceptions of Selected Drop-outs Who Had the Ability to Complete High School If They Had Persisted.

This study was designed to develop hypotheses for guidance, curriculum, and administrative practice in the schools as well as further studies of the drop out. A selected sample of six subjects was drawn on the following criteria: Male, white, age 16-17, I.Q. range 90-100 and sufficient test evidence and past academic performance to indicate that they would have finished the high school if they had persisted. A sample of 50 students meeting the same criteria as above but still attending the same high schools (persisters) was drawn and tested for comparative purposes.

The 6 subjects, all high school dropouts, were administered the following instruments: WAIS, Modified Semantic Differential, Modified Engel Self Concept Scale, Self Anchoring Scale and a ten hour interview. The interview was developed along the following referents, Curriculum, Teachers, Other School Personnel, Friends, Self, Parents, Home, Siblings, the Future.

The comparative group of persisters was administered the instruments with the exception of the WAIS and the interview. Test data between the dropouts and the persisters were compared. The bulk of the findings consists of a thematic analysis, longitudinal and cross sectionally of the interview data of the dropouts and the hypotheses for further explorations. The investigators found that 1) The perceptions of the referents and subsequent withdrawal from school were idiosyncratic to each subject; 2) Each subject gave a number of indications that the school was an unsatisfying experience, although these were subtle and not overt actions in most cases; 3) There was a complete failure of school to attempt to interdict any of the symptoms at any stage that indicated the student was going to terminate his schooling; 4) The problems of the subjects in the school were now reflected in their adjustment to the world of work; 5) The test instruments with one exception failed to discriminate between the persisters and dropouts.

Robert H. Fenske, University of Minnesota. Association Between Local College Availability and Plans for College Attendance of Public High School Seniors with Differing Attributes and Socio-economic Characteristics.

This study primarily investigated the association between the presence or absence of an institution of higher education in the local community and the plans of public high school graduates to attend college. Analysis focused on the extent to which this association was conditioned by differences in the following characteristics of the graduates and their socio-economic backgrounds: a) Scholastic ability, b) Parents' educational attainment, c) Type of father's occupation. These background characteristics were held constant to clarify the association of prime interest. Such an analytic procedure had evidently not been utilized before on this problem important to planning for higher education. Ancillary analyses were also made of closely related topics within the available data. Subjects were nearly all of the graduating seniors (N = 3952) in ten urban Wisconsin communities - four of which had university extension centers; two, state universities; and four, no colleges. The main method of analysis was elaboration by partials. The strength of association was measured with gamma.

Salient findings: The presence or absence of a local college was critically important to college attendance plans for comparatively few of these graduating seniors. Scholastic ability and socio-economic background characteristics usually were the important factors in such decisions for the great majority. Interesting and significant variations in association between plans and local college availability occurred, however, with differing combinations of these characteristics. Such variations were markedly different between boys and girls.

Striking differences existed among the communities in distribution of the graduates' ability and background characteristics, even though these communities were similar in many basic demographic and economic factors. Such differences were markedly more determinative of college attendance plans than local college availability. It was revealed, however, that for some graduates, --e.g., those of average ability and average to low socio-economic status (particularly girls), local presence of a college is of critical importance in their educational plans. This study underscores the pervasive importance of personal factors (e.g. scholastic ability and family influence) in decisions about college attendance and reveals that by comparison, local availability of a college is relatively uninfluential upon such decisions.

J. Kenneth Little, University of Wisconsin. Occupations of Non-College Youth.

One-third of the 1957 male graduates of Wisconsin high schools were traced to learn of their occupational status eight years later (1964-65). Information was received from 85% of the sample (N = 4,185). Extensive background information about these youth had been collected shortly before their graduation. The new information covered their post-high school educational and occupational history. The findings show a large overlap in occupational status among youth with differing amounts and types of education. Occupational attainment like educational progress appears to depend upon: a) personal qualities such as general ability and habits of achievement, and b) cultural level of family and community. A comparison of stated occupational aspirations with later occupational attainments shows that most non-college youth were realistic in assessing their probable occupational opportunities and achievements. Implications for the education and non-college bound youth are drawn.

SESSION 49

Paper session --STATISTICAL METHOD AND THEORY - I

R.O. Collier, University of Minnesota and F.B. Baker, University of Wisconsin. Factors Affecting the Agreement of the F-Test Under Permutation and Normal Theory.

Scheffé (Chapter 9, 1959) presents considerable evidence which indicates both the size and power of the F-test under permutation and under normal theory agree. Because of this agreement, the present paper systematically investigates three factors: Heterogeneity of block variance, block-treatment interaction, and kurtosis of the basal response data, which should according to statistical theory affect the agreement. Size and power data yielded by two simple randomized block designs -- a 5 block by 3 treatment and a 12 block by 3 treatment - were used to study the first and second factor. The size and power data from a single randomized block design - a 3 block by 3 treatment with six observations per cell - was used to investigate the first and third factors. The measure used in the analysis was the discrepancy between empirical size and a normal theory value of .05 and the discrepancy between empirical power and normal theory power of .60 at a size of .05.

The empirical results indicated that agreement of the size of the "F-test" under permutation and under normal theory is affected by block-variance heterogeneity. The agreement of power of the F-test under the two models was immune to this factor. Power was, however, affected by the interaction of number of blocks and level of block-treatment interaction. Empirical power decreases markedly as block-treatment interaction increased in the five block design but not at all in the 12 block design, hence the interaction. Despite the highly kurtoic data employed, γ_2 = 3.0, 5.0, 7.0, the agreement of both size and power was unaffected by this factor in the present study.

The insensitivity of the F-test under permutation to the factors involved in the present study suggests the permutation procedure should be more widely used by educational researchers.

Peter Dunn-Rankin, University of Hawaii and Frank Wilcoxon, Florida State University. Approximations to the Range of Rank Totals in the Two-Way Classification.

In addition to determining whether or not several items belong to the same population, an experimenter may also be concerned with significant differences between pairs of items or treatments. In making nonparametric multiple comparisons, the concept of the range has been utilized in the two-way classification.

This paper presents the results of an investigation into the probability distributions of the range of rank totals and gives a procedure for generating an approximation to the true distribution. A comparison of the results of this approximation with an extensive criterion of computer generated true and sample distributions, and with other approximations supports the accuracy of this method. Accurate estimates of the critical ranges necessary to reach significance at three commonly used alpha levels, where the number of judges and items are less than or equal to fifteen are illustrated in tubular form. An algorithim for finding critical ranges for more than fifteen judges is also shown.

The table produced within the framework of this investigation should aid in the scaling of stimuli and contribute to making tests of significance in any two-way classification analysis of variance by ranks. The sample and true distributions that have been generated should provide a firm base for validating other theoretical approximations to the true distributions of the range of rank totals. The approximation formulated in this study also appears to have substantial value for studies of the range of rank totals in the one-way classification, i.e., a classification of treatments by replications.

Robert James Schmeidler and Richard H. Lindeman, Columbia University. *Some Procedures for Analysis of Incomplete Designs.*

Incomplete designs in educational experimentation are useful because satisfactory inferences can often be drawn on the basis of relatively few observations. Sometimes such designs are unavoidable; not all combinations of conditions can be obtained. Similar "designs" can also occur by accident, e.g., if a number of observations are missing. This paper discusses procedures for analysis in two cases in which data are incomplete for one or both of these reasons.

An example of the first case considered is that in which each member of a team of judges rates all individuals in a small group. In this case, the data can be arranged in the form of a two-way analysis of variance with one observation per cell. If any of the cells are empty, then estimates of row and column means (obtained through inversion of the normal equations) are unbiased but dependent.

The second case is that in which each of several teams of judges observes one of several groups. Each judge in a given team rates all of the subjects in one group and only those subjects. The team-group pairs are assumed to be independent of one another and the analysis is performed on the pooled experiment. Any or all of the sub-experiment may have missing observations if all judges in a team do not rate all subjects in the group.

Procedures for inference in these cases are discussed and examples are given. The ordinary analysis of variance tests of significance cannot be used. An alternative, applicable in any problem with several hypotheses is suggested.

Andrew C. Porter, University of Wisconsin. *A Chi-Square Approach to Discrimination Among Occupations, Using and Interest Inventory.*

The study proposed the use of a chi-square technique for dealing with the problem of weighting items which are used for discrimination. An empirical comparison was made of the chi-square technique with the technique presently in use for the Kuder Vocational Interest Inventory Form D. The instrument used for purposes of the investigation was the 100 triadic items of the Kuder Vocational Interest Inventory Form D. The main hypothesis tested was: A set of weights, established by the chi-square technique, discriminates among more than two similar occupations better than the weights used by the Kuder. The study was based on a sample size of 1902 male responses to the 100 items of the Kuder Test, and made use of a 1604 Control Data Corporation Computer to facilitate the computation involved with large samples.

Empirical chi-square weights were derived for each of five occupations considered by observation to be similar in nature. The derived chi-square weights were applied to cross-validation groups, and percentages of correctly classified males were determined for each of the occupations. The criterion for correct classification was the known occupation. The resulting percentages of correctly classified males were better than what would have been expected by chance for each of the five occupations. Double cross-validation studies of the derived chi-square weights indicated that the weights were general in nature and not just indicative of the sample upon which the study was based.

The same data scored by chi-square weights was also scored by Kuder weights, and the percentages of correctly classified males were again determined. An occupation by occupation comparison of the percentages of correctly classified males of the chi-square and the Kuder techniques indicated a significant increase in favor of the chi-square technique for two of the five occupations and no significant difference for the other three occupations.

Lawrence Wightman,, Cornell University and Ray M. Snider, New York State University College at New Paltz. Observer Reliability in Interaction Analysis.

The Scott Index is a measure of the reliability between two Flander's Interaction Analysis (I.A.) codings of the same verbal behavior, be it inter-observer reliability or intra-observer reliability. I.A. observers have relied on the Scott almost totally for indications of their reliability and, by inference, their validity.

This paper explores relationships between the Scott Index and the Darwin chi-square comparison of pairs of I.A. matrices. By looking at these relationships over almost one hundred matrix pairs not representing the same verbal behavior, over matrix pairs representing reliability data, and over matrix pairs in which the data have been systematically manipulated, certain weaknesses show up in the Scott Index.

The authors show results of the Darwin chi-square comparisons in a parallel manner and contend that the Darwin comparison is a more valid indication of reliability.

SESSION 52

Paper session --RESEARCH ON TEST-TAKING BEHAVIOR

Marion Turpan, Fairleigh Dickinson University and Gerhard Lang, Board of Education of the City of New York. *Test Performance as a Function of the Order of Item Difficulty in Multiple-Choice Tests.*

The following hypothesis was tested: Students administered a multiple achievement test constructed in an ascending order of difficulty will achieve higher test scores than those students given a test of either randomized order or of descending order of difficulty.

The subjects consisted of 192 freshmen, 150 males and 42 females. They represented a cross-section of the various colleges of the Fairleigh Dickinson University.

It was concluded that students administered a multiple-choice achievement test constructed in an ascending order of difficulty did not achieve higher test scores than those students given a test of either randomized order or of descending order of difficulty. In this study, therefore, test performance appears to be unrelated to the arrangement of items in order of difficulty.

William L. Goodwin, University of Wisconsin. *The Effects on Achievement Test Results of Varying Conditions and Experimental Atmosphere, Notice of Test, Test Administration, and Test Scoring.*

The researcher is often faced with selecting that means of administering the tests to maximize the validity and generalizability of any conclusions reached. Four null hypotheses were tested to determine the differential effects of: experimental atmosphere and absence of same; notice of test (10 school days) and no notice (one school day); teacher administration and outside administration of the test; and teacher scoring and outside scoring of the test.

The experimental unit was the classroom. Sixty-four sixth-grade classes, each from a different school in a large midwestern city, were ranked and grouped into four strata on the basis of previous arithmetic achievement. Within each strata, classes were randomly assigned to one of the 16 experimental treatments generated by a 2^4 factorial design using the four independent variables as listed above and in connection with a recent arithmetic achievement test as a response measure. Experimental atmosphere was created using written instructions, and test notice was given by mail. The outside test administrators and scorers were graduate and undergraduate students. Resulting class means of the 64 classes for each of the three sub-tests in the exam were subjected to a 4×2^4 analysis of variance. The error term was composed by pooling selected higher-order interactions.

Tests of the main effects revealed significantly higher class means on one of the three sub-tests for those classes receiving 10 school days' notice of the upcoming test and significantly higher class means on all three sub-tests for those classes whose regular teacher administered the test. In addition, several two-factor interactions were significant.

Clinton I. Chase, Indiana University. *Changing Response Habits Through Changes in Test Content.*

Can the tendency to be agreeable on agree-disagree test items be a) increased, and b) decreased by alterations in test content?

Two forms of an agree-disagree type test were built. The first 26 and the last 26 items taken directly from a personality test were identical between Form A and B. Between these two groups of items, Form A had 18 items that consistently drew disagree responses. Thirty-five college undergraduates took Form A; thirty-four took Form B.

For analysis, the first 26 items were called the pretest, the next items response training, and the last 26 items the posttest. Papers were scored for "agree" responses only.

Form A's posttest had significantly more "agrees" than the pretest. Form B's posttest did not differ from the pretest on "agree" responses.

Groups were then divided into "high agrees", Low agrees" on pretest scores. Pretest high agrees did not significantly change performance on the posttest on either form; however, pretest low agrees significantly increased their number of "agrees" on the posttest for Form A; significantly reduced their number of "agrees" on the posttest on Form B. Hypotheses are submitted to explain these findings.

For persons with initially a "high" response preference, test content manipulation does not appear to alter that preference; response mode for persons with low initial preference may be altered by test content manipulation.

Gary L. Marco, Educational Testing Service. Validation of a Measure of Specific Test Anxiety.

The purpose of this study was to validate the Anxiety Differential - an instrument developed by Husek and Alexander (1963) to measure specific test anxiety. It was predicted that scores on the Anxiety Differential in a testing situation would be higher than scores in a non-testing situation. Further, it was predicted that scores on the Anxiety Differential in a testing situation designed to arouse a moderate degree of anxiety would be higher than scores in a testing situation designed to arouse minimal anxiety.

A version of the Anxiety Differential consisting of 14 concept-adjective semantic differential scales selected from anxiety-relevant scales reported by Husek and Alexander and eight filler items was composed. The instrument was administered during a regular class period and at the beginning of each of four 50-minute test periods to 154 college seniors enrolled in educational psychology. Three of the tests were in educational psychology and one was in logic. Since scores on the latter test were not used to determine course grades, it was expected that the tests in educational psychology would arouse more anxiety.

Results confirmed the predictions. All of the means in the testing situations were significantly higher than the mean in the non-testing situation. And the mean of the Anxiety Differential administered just prior to the logic test was lower than means of the Anxiety Differentials administered just prior to the educational psychology tests. In addition to these results, discrimination indices for the individual scales of the Anxiety Differential are reported.

SESSION 54

Paper session - PROBLEM SOLVING AND CREATIVITY.

Evan Keislar and Carolyn Stern, University of California. The Value of Different Problem Solving Strategies with Young Children as Related to Mental Age.

Three experimental studies were conducted to test the following hypotheses: 1) Young children who are taught a strategy for successive testing of hypotheses in concept identification problems will be superior in this task to children who are given practice on problems without training; 2) Children who are taught to test several hypotheses at once will be superior to children who are taught to test hypotheses singly; 3) Mental ability level is more closely related to learning the Multiple Hypothesis strategy than the Single Hypothesis strategy; 4) There is an interaction between mental ability and performance under the two strategy conditions.

All Ss were third-graders within the age range 7-3 and 8-8. Experiment I used 110 subjects; experiment II used 107 subjects. In experiment III, 132 children were used with a high average ability. The three experiments used the same concept identification problems involving identifying the basis for matching figures (number, size, color, or shape). Children were taught and tested by means of programs using projected slides and taped commentary. The instructional programs were improved following each experiment. Using a confidence level of .05 throughout, hypotheses 1 and 3 were supported in all 3 experiments. Hypothesis 2 was revised in experiments 1 and 2, where less capable populations were used, suggesting that the Multiple Hypothesis strategy was very difficult for this population. With the improved program used in Experiment 3, both hypotheses 2 and 4 were supported.

John P. Casey, Southern Illinois University. An Experimental Study of Creativity and the Solving of Social Problems.

The purpose of the study was to present empirical evidence of the difference or lack of difference between more creative and less creative high school students in their methods of solving social problems.

Two groups of senior high school students, 18 in each group, were identified as more creative and less creative by means of Getzel's and Jackson's Creativity Tests, and matched on the basis of their scores on the California Mental Maturity Test, participated in the experiment. Three masked situational problems of a social nature were administered to the subjects. Spearman's rank-order method for finding correlations; a t test for correlated samples, and chi square were used to test the null hypotheses.

The findings were that: 1) There was a relationship between the two groups in their initial choices and acceptance of the bits; 2) There was no significant difference between the 2 groups in the number of bits selected; 3) There was no significant difference between the 2 groups in the presented solutions and generalizations selected.

SESSION 54

George D. Yonge, University of California at Davis. *Effect of the Number of Unusual Uses on Functional Fixedness.*

Flavell presented evidence in support of the hypothesis that there is a progressive decrease in functional fixedness as the number of unusual functions experienced in the pre-problem training period increases. His results are ambiguous because the *position* of the experience of the usual function in the training period covaried with the *number* of unusual uses experienced. In the present study, the number of novel functions experienced was held constant while the position of the usual function (switch) in the series of training experiences was varied.

The Maier two-string problem was used to assess functional fixedness. 115 female university students were assigned to one of five groups. The control group provided the baseline necessary for demonstrating functional fixedness. Group E_1, the conventional experimental group, demonstrated the comparability of the present procedure to previous studies. Group E_2, E_3, and E_4 experienced the switch in its usual and four unusual functions during the training session. E_2 experienced the switch function first, E_3 experienced it in the middle position, and E_4 experienced it in the last position.

According to Flavell's hypothesis, groups E_2, E_3, and E_4 should not differ (the number of unusual uses experienced was constant). An analysis of variance ($p < .02$) was followed by Duncan's test. Contrary to Flavell's hypothesis, group E differed ($p < .05$) from E_3 and E_4. Thus, when the usual function was experienced *first,* functional fixedness was manifested even though four uncommon functions were subsequently experienced. A phenomenological interpretation of the results was given.

Sara W. Lundsteen, University of California at Santa Barbara. *Abstract, Functional and Concrete Levels of Thinking in Children.*

To further dimensionalize variables of abstract, functional, and concrete levels of thinking, a 42 item test of word meaning on vocabulary was revised and administered to 178 pupils in grades 3 and 6. In grade 3, pupils continued preference for concrete responses; in grade 6, pupils favored the abstract responses (.01). In accord with theory from Piaget revision of the functional category appeared to restore performance to a shared level rather than to a level confounding the abstract category. Again, when children were classified according to dominant response, (at least 40% of the choices in that area), it was found that a majority of cases in grade 3 could be assigned to concrete style (13%) and in grade 6, to abstract style (67%). Between the experimental test and tests of reading, critical listening, achievement, and problem-solving, *r*'s were .66, .67, .67, .72 respectively.

SESSION 54

Richard E. Owens, Kansas State University. The Relationship of Creative Thinking Ability to Extreme Over and Underachievement.

The purpose of the study was to investigate how creative thinking ability related to over and underachievement at the intermediate level of the elementary school. The sample was selected on the basis of a comparison of intelligence and achievement scores on standardized group tests. The subjects were then tested with a complete battery of creativity tests including group nonverbal tasks, group verbal tasks, and individual, oral verbal tasks.

The overachievers were found to score significantly higher than the underachievers on flexibility, originality, and adequacy of response. However, there was no significant difference between the groups on fluency, elaboration, activity, and total creativity. The overachievers were found to be significantly higher than the underachievers in verbal creativity on all subareas and the total. However, there was no significant difference in the two groups on nonverbal creativity.

Both groups made significant gains in going from the written, group administration to the individual, oral administration, regardless of the order of the tests. The overachievers consistently gained more but the difference was not significantly large.

The correlation between total intelligence and creativity was low positive and statistically non-significant. Achievement correlated low, but significantly, with total creativity. Factor analysis of the basic components of intelligence, achievement, verbal creativity, and nonverbal creativity yielded four distinct factors for the sample as a whole.

There were extreme individual differences in the creativity scores of both groups. Many implications were made to teaching methods and curriculum at the intermediate level of education.

SESSION 55

Paper session --TEACHER BEHAVIOR AND PUPIL GROWTH

Robert S. Soar, Temple University. *Teacher-Pupil Behavior and Pupil Growth.*

This research is an empirical study of the relationships between teacher-pupil classroom behavior and pupil growth.

Procedure: The Vocabulary, Reading and Arithmetic sub-tests of the *Iowa Tests of Basic Skills*, verbal and non-verbal tasks of the *Minnesota Tests of Creative Thinking*, the *Children's Manifest Anxiety Scale*, and the *Dependence-Proneness* scale were administered in the fall and spring in 56 elementary classrooms, grades three through six. In addition, the *My Class* inventory and the *Pupil Survey* (Self-Initiated Work only) were administered in the spring. During the year, teacher-pupil behavior in these classrooms was observed using two observation schedules --the *Interaction Analysis* and another schedule made up of items from several revisions of the *Observation Schedule and Record* and the *Hostility-Affection Schedule*.

Residual gain scores were calculated for the pupils by first estimating true gain by a procedure outlined by Lord, and the relationship with initial standing was eliminated. The classroom observation data were reduced by two factor analyses of approximately 60 measures derived from the observation schedules, and factor scores were calculated for each teacher on each factor. These factor scores were correlated with the residual gain means for each classroom to determine whether relationships exist between the nature of the classroom and pupil growth.

William S. LaShier, Jr., University of Kansas. *An Analysis of Certain Aspects of the Verbal Behavior of Student Teachers of Eighth Grade Students Participating in a BSCS Laboratory Block.*

A study was conducted to determine the relationship between the verbal behavior of student teachers and the achievement and constructive attitudes of eighth grade students during a six week period of instruction devoted to a study of a BSCS unit entitled *Animal Behavior*. Ten student teachers taught the BSCS Laboratory block to 239 eighth grade students in six junior high schools. The student teachers were observed once a week during the six week laboratory block term by one of three observers using the system of interaction analysis developed by Flanders. An I/D Ratio was established for each student teacher.

A pre- and post-test design using the Animal Behavior Test provided a measure of the gain in achievement. Information concerning the attitudes of the students toward their student teachers and their school work was obtained from the Michigan Student Questionnaire. The scores of the students on the California Achievement Tests in Reading were also obtained.

The findings of the study indicated a significant positive relationship between gains in median achievement on the Animal Behavior Test and the I/D Ratio of the student teachers. There was also a significant relationship between the I/D Ratio of the student teacher and the class median on the Michigan Student Questionnaire.

SESSION 55

June J. Slobodian, Livonia Public Schools and O. L. Davis, Jr., Kent State University. *An Analysis of the Relationship between Teacher Behavior and Pupil Achievement in Beginning Reading.*

This study explored 238 first-grade children's perceptions of interactions in the reading group situation to discover whether first grade children saw teachers discriminating against boys. Observations of the interactions in the actual reading group situation were then gathered to determine if teachers in fact differentially responded to boys and girls. Achievement scores of boys and girls in ten first grade classrooms of the study were then compared. Children's perceptions were obtained by use of questionnaire-interviews. Observers were trained in using a new Reading Observation Record. developed for the study in order to record the classroom interaction units. Achievement data was collected from a standardized test administered to the children at the end of first grade. Obtained data was treated by use of *chi* square and analysis of variance (Lindquist, Type I).

Results indicated that children do perceive differential treatment of boys and girls by first grade teachers and that boys were seen as being discriminated against in the reading situation. However, observation of teacher-pupil interaction in reading groups revealed that teachers: (1) did not give boys less opportunity to read and respond, and (2) did not praise nor criticize boys more than girls. Achievement of boys and girls, contrary to expectations, did not differ significantly. Discussion of these findings is set within the framework of theory regarding instruction in beginning reading and teacher influence in the classroom.

John M. Kean, University of Wisconsin. *An Exploration of the Linguistic Structure of Second - and Fifth-Grade Teachers' Oral Classroom Language.*

The objective of this study was to describe the linguistic structure of elementary school teachers' classroom language as a first step toward the study of the interaction between teachers' language and children's language in the classroom. The language of 10 second-grade and 10 fifth-grade teachers was recorded during 5 forty-minute periods of regular classroom operation. Linguistic data including amounts of language (communication units and mazes), structural patterns, variable components of the structural patterns (e.g. kinds of nominals used as subjects), degree of subordination, and vocabulary diversity were obtained. Descriptive statistics are reported for all measures. Intercorrelations of major criterion measures were computed. The basic description of teachers' language indicated that there does not appear to be much difference in the structure of the teachers' language or in the diversity of vocabulary at second-and fifth-grade levels. However, the correlational analysis revealed relationships between the variables that would seem indicative of grade level differences. The relationship of the results of this study to other studies of teaching and to the teaching of language arts are discussed.

SESSION 57

Paper session: --STUDIES OF STUDENT CHARACTERISTICS

J. Robert Staffieri, Pineland Hospital and Training Centre, Pownal, Maine and Boyd R. McCandless, Indiana University. *A Study of Social Stereotype of Body Image in Children.*

The purpose of this study was to investigate the role of body image stereotypes in children with reference to development, interpersonal functions, and social functions.

Ss were 90 male children from 6 to 10 years of age. The sample population consisted of 18 Ss in each age level who were grouped (according to their relative fatness, muscularity, and thinness) on the basis of the Ponderal Index and teacher rating of body type. Each S was assigned to assign 39 adjectives (which were descriptions of various personality/behavior traits) to silhouettes which represented extreme endomorph, mesomorph, and ectomorph body types. Sociometric data were obtained from all classes (grades 1 through 5) for the number of friendship choices (acceptances) and number of dislike choices (rejections) received by each S in the study.

In part, the results of the study indicate that: 1) Boys from 6 to 10 years old have a common concept of behavior/personality traits that are associated with various body types; 2) All the significant adjectives assigned to the mesomorph image were favorable (16); the adjectives assigned to the endomorph were unfavorable (socially) and primarily socially aggressive; the adjectives assigned to the ectomortph were primarily unfavorable (personally) and of a generally socially submissive type. 3) Ss showed a clear preference to look like the mesomorph image. 4) Mesomorph Ss received a consistently high number of acceptance choices and endomorph Ss received a consistently low number of acceptance choices; 5) There was no difference in the number of rejection choices received by ectomorph, mesomorph, and endomorph Ss.

Gerald T. Kowitz, University of Houston and Norma G. Kowitz. *School Attendance as an Index of Guidance Needs.*

A high correlation is usually assumed between school attendance and achievement. Attendance is also believed to be a function of the child's adjustment to the school environment. If these relationships can be substantiated, attendance data should become a powerful screening device for the guidance services.

The attendance records of 235 second grade students were studied. Each child was classified by his teacher either into one of two problem groups (Learning-Educational or Personal-Social) or into one of two non-problem groups (Enrichment or Average-Normal). The teachers also represented one of four groups: Either being from a school serving a higher or lower economic area and as having extended or limited classroom experience.

Three types of absences were studied: Single day, multiple day and total days absent. Single day absences were expected to be most closely related to guidance needs. Multiple classification analysis of variance revealed no significant differences in the average number of days absent among the four pupil groups, the four teacher groups or among their interactions, regardless of whether the data were single day, multiple day or total days absent.

Whether the data were studied in terms of pupil or teacher groups, the average number of days absent was never significantly different from zero.

No support was found for the idea that school attendance is an index of pupil achievement or adjustment.

John A. Finger, Jr., Rhode Island College, Predictors of Change in School Performance at Entrance to the Junior High School.

The purpose of this study was to determine whether the drop in grades which is reported to occur as students move from elementary school to junior high school could be related to: intelligence, academic motivation, aspirations, or other student characteristics.

Sixth-grade teachers were interviewed to obtain information concerning their students' academic motivation, socio-economic status, dependency, classroom and social behavior. Teachers were asked to predict which students would do less well in junior high school. Students were tested with the Personal Values Inventory on entering the junior high school. California Mental Maturity and California Achievement test scores were obtained at the sixth grade level from a state-wide testing program.

A factor analysis of the achievement, intelligence and academic motivation data identified four factors: intelligence-achievement, academic motivation, academic plans, and youth culture involvement. These factors were related to marks achieved in the elementary school, but the change in performance at the junior high school was related to academic motivation, academic plans, and youth culture involvement but not intelligence achievement. The discriminant function using the variables which defined the factors was utilized to assign students to groups of high or low elementary marks and high or low junior high school marks. Students correctly assigned by the discriminants were compared with those incorrectly assigned to determine whether the teacher interview data could account for the discrepancies.

The characteristics which cause students to drop in performance in junior high school are present at least as early as the later years of elementary school.

H.B. Rose, Waukesha Public Schools, Waukesha, Wisconsin. An Acceleration Program for Superior Students in the Waukesha Elementary Schools.

In the Waukesha Public School System, we have an early admissions program at the kindergarten level. However, it was felt that this did not offer sufficient flexibility and that provisions should be made for proven superior students to be accelerated one grade.

After studying various plans, it was decided to accelerate a selected group of students after the completion of the second grade. The students who are selected, and after the parents approve participation, attend summer school for six weeks to cover the essentials of third grade. After successful completion of the summer school program, the students are placed into the fourth grade in their district schools.

The Presentation will discuss: 1) Screening Procedure and steps and criteria used; 2) Selection Procedure - Selection Committee; 3) Parent

Orientation Meeting; 4) Outline of summer school program; Followup studies for all accelerated groups and comparison groups.

Leonard A. Marascuilo and Joel R. Levin, University of California at Berkeley. *Inter- and Intra-racial Group Differences in the Perception of a Social Situation.*

The hypothesis studied is that one's perception of the extent to which inter-racial socialization has been achieved is influenced by his particular racial membership, as well as his personal relationships with members of other races. Data were obtained from questionnaires completed by all 7th, 8th, and 9th grade students in 3 Junior High Schools of a northern urban community. The questionnaire was designed to measure students' attitudes toward a major reorganization program, the purpose of which was to produce an improved racial balance within the schools. The responses were obtained one year after the initiation of the new school program. The students were also requested to indicate the amount of inter-racial association in which they and their classmates participated. The latter varied greatly with the race of the respondent. It was found that members of the same race tended to answer a given question in the same way. At the same time, however, it was apparent that individuals who had made friends from other racial groups, responded differently to the questions concerning inter-racial mixing than did members of their particular race who had not formed inter-racial friendships. These results suggest that one's racial membership *per se*, and his individual relationships with persons of other races contribute to differences in his perception and attitudes concerning inter-racial socialization.

Barbara H. Long, Goucher College and Robert C. Ziller and Edmund H. Henderson, University of Delaware. *Developmental Changes in the Self-Concept of Adolescents.*

This study investigated the self-concept in relation to age/grade and sex among 420 students (30 boys and 30 girls of proper age-for-grade in each grade, 6-12). The Self-Social Symbols Tasks, a non-verbal test of various (7) components of the self in relation to others, was administered to all subjects. Split-half reliability coefficients ranged from +.58 to +.92 with a median of +.82 for the 12 measures. Intercorrelations showed the 7 components to be largely independent. Results: (A) *Self-esteem* increased with grade level ($p= .05$). (B) *Dependency.* Social dependency increased until 9th grade, declined thereafter ($p =.01$). (C) *Power.* Power of the self in relation to father declined over grade level ($p= .05$), and was less than power in relation to teacher or principal ($p = .001$). Senior high students showed a more egalitarian relationship to teacher and principal than did those younger ($p =.025$). (D) *Identification.* 1. Boys included more others in the self category than did girls ($p = .001$. 2. Younger girls (grades 6, 7, 8) identified less with mother than did older girls ($p = .001$). 3. 12th grade boys identified less with teacher ($p = .05$), father ($p = .05$) and friend ($p = .10$) than did younger boys. 4. Among the girls, identification with friend increased with grade level ($p = .05$); an opposite effect was found among the boys after grade 9 ($p = .10$). f. In all grades but the 12th, boys identified more with father than did girls ($p = .01$).

SESSION 58

Paper session --TEST DEVELOPMENT

T. R. Husek, University of California, Los Angeles. Different Kinds of Evaluation and Their Implications for Test Development.

A test may be useful for some purposes and not for others, and current treatment of validity and reliability stress that appropriate evidence is dependent on the context of the use to which the test is put. However, in the field of evaluation, most students are instructed as if there were only one purpose for a test, to maximally discriminate among testees.

A good evaluation instrument provides information useful in making the desired evaluation, such as the evaluation of students at the end of a course, the evaluation of student change, or the evaluation of a course itself. Different ways of choosing items arise as a function of the different kinds of evaluation desired, and conflicting item choice rules flow from the different kinds of evaluation desired.

This paper will also suggest that teachers should not use the item analysis procedures they are ordinarily taught. Part of this argument will suggest that the teacher who consistently uses a simple item analysis procedure for improving tests will gradually develop a mental ability test rather than a good achievement test for a course.

Lastly, it will be suggested that item sampling can be used for part of a test to evaluate a course while the rest of the test is used to evaluate the students, and that sequential testing within a single test can be used in the schools to permit more information to be obtained at important decision points.

Arieh Lewy and Christine McGuire, University of Illinois, College of Medicine. A Study of Alternative Approaches in Estimating the Reliability of Unconventional Tests.

Despite the availability of sophisticated methods of estimating reliability, many test producers and most test consumers are still content with a single estimate of test reliability based on a unidimensional concept, e. g. internal consistency, even when inherent characteristics of the instrument violate explicit assumptions imposed by the technique employed.

This paper describes a research study of the reliability of selected written simulation exercises that differ from conventional tests in the following respects: 1) items are differentially weighted; 2) items are interdependent; 3) no examinee responds to all sections of an exercise; and 4) differential amounts of feedback are available to different examinees depending on their own decisions at various points in an exercise.

In this study, the problem of estimating reliability was investigated from the point of view of the "generalizability" (Cattell, Cronbach) of responses to the exercises. Utilizing variations of this concept, four different approaches to the measurement of reliability were employed, and the estimate yielded by each was analyzed in terms of the universe to which results were generalizable. Results of the four methods are presented, the assumptions underlying each are analyzed and the significance of each discussed.

SESSION 58

Jack R. Frymier, Chio State University. Measuring Creativity and Delinquency-Proneness With a Picture Preference Scale.

This paper describes two studies: one involved creative and non-creative medical students, and the other involved delinquent and non-delinquent adolescents. Both studies compared the responses of known groups to The Chio State Picture Preference Scale, a 100-item, objective, non-verbal measure of personality and motivation currently under development. Inferences from these comparisons are made.

Study One: All of the freshmen (N = 147) in the Ohio State University College of Medicine were tested with the picture preference scale. Using four external criterion measures (Creativity Personality Scale of the OAIS, Complexity and Autonomy scores of the Omnibus Personality Inventory, and the Dogmatism Scale) high-creative and low-creative groups were identified. Comparisons of these students' responses to the picture preference scale were undertaken by item analysis.

Study Two: All of the male delinquents (N = 107) committed to the Ohio Juvenile Diagnostic Center for assignment to an incarcerating institution during a two-month period were tested with the picture preference scale. Three groups of non-delinquents meeting various criterion measures of academic motivation (teacher judgment, scores on a motivation index, and over- and under-achievement scores) were selected for comparison purposes. A study of the responses of delinquents and non-delinquents to the picture preference scale was undertaken by means of item analysis.

Kenneth Urial Gutsch, University of Southern Mississippi. Objective Measurement in Instrumental Music Performance.

This study sought (1) to determine if an objective measurement of instrumental music could be obtained on sight-reading rhythms, and (2) to differentiate degrees of attainment in instrumental music achievement while sight-reading rhythms.

Through graphing concepts introduced by Joseph Schillinger, mathematical constructs were used to develop equivalent forms of an instrumental music performance test which was then individually administered. Complete test protocols for 771 subjects from grades 5 - 12 revealed a high degree of consistency for the two forms of the test with the coefficient of correlation for test reliability falling at .95. Randomly selected tape recordings of testing sessions reflected high scorer reliability with a score-rescore coefficient of .99.

Multiple and partial correlations revealed the influence of such variables as age, instrumental experience, grade level in school, scholastic average, and IQ upon performance, and t-tests results indicated differences between any two levels of experience when experience was defined in years.

Results of the study indicate that within the confines of the methods and techniques introduced during the study, the test has both test and scorer reliability and that of those variables studied, instrumental music experience was most influential as a final determinant of test score.

SESSION 60

Paper session --MODIFYING TEACHER BEHAVIOR

Bert Y. Kersh, Oregon State System of Higher Education. Classroom Simulation: Further Studies on Dimensions of Realism.

Findings from preliminary research with a simulated classroom used to train teachers indicated that student teachers actually may learn better when the simulation is unrealistic, so long as they first have simulated experiences under realistic conditions. Since then, three studies have been completed which (1) tested the relative importance of the realistic-unrealistic experience in the simulated classroom, (2) explored response and feedback dimensions of realism in simulation, and (3) explored transfer effects of simulation experiences to actual classroom situations. Findings from the first experiment failed to substantiate the preliminary findings regarding the order of realistic-unrealistic experience. Gains in performance were substantial regardless of variations in screen size. The second experiment indicated that mode of response interacts significantly with pretest performance ratings. That is, student teachers who scored below the mean on the pretest performed better on the post-test after "acting out" as contrasted with "describing" their responses to simulated problems; whereas, students scoring above the mean on the pretest gained as much after simply describing their responses as they did after enacting them. Finally, in a third study, supervisor ratings of student teachers in actual student teaching situations revealed that those who had completed simulation instruction were able to assume full responsibility for the classroom approximately three weeks before those who had no experience in the simulated classroom.

Paul A. Twelker, Oregon State System of Higher Education. Prompting as an Instructional Variable in Classroom Simulation.

Typically, classroom simulation training, as developed by Kersh, relies upon a "learning by discovery" method to teach a student teacher modes of operation within certain behavioral limitations in response to a variety of problematic situations. Studies on discovery learning generally indicate that prompting subjects with principles increases transfer, reduces the time required for learning, and increases effectivity toward the learning experience. This experiment investigates the value, in terms of increased transfer and efficiency, of prompting subjects with information relevant to (1) how to respond, and (2) to what to respond during a simulation experience.

Seventy students selected from elementary education classes were randomly assigned to each of four treatment groups and a control group. A 2 x 2 factorial design was used to study the two types of prompts. The four treatment groups were: (1) problem prompt given; response prompt given, (2) problem prompt given; response prompt not given, (3) problem prompt not given; response prompt given, (4) problem prompt not given; response prompt not given. A control group was used to assess maturational factors during training. These subjects were not given the learning experience.

Subjects were tested immediately on a test of transfer. Measures of affectivity, time to learn, and learning program error rate were obtained.

Evidence suggests that learning and transfer are enhanced when students are prompted with the necessary principles to recognize simulated problems and to respond to them. Withholding these principles reduces performance, and wastes time. Generally, the results support a prompting procedure over a feedback procedure for teaching college students to solve the problems similar to those presented in this study. The results point the way to the eventual adoption of more efficient methods of simulation training than are currently available.

Frederick J. McDonald, Dwight W. Allen, and Michael E. J. Orme, Stanford University. Effect of Self Evaluation and Social Reinforcement on the Acquisition of a Teaching Behavior.

The independent variable in this study was the mode of reinforcement applied to acquiring a teaching behavior. Three modes were used: (1) subjects evaluated their own behavior by viewing their videotaped teaching performance; (2) an experimenter reinforced verbally each enactment of the desired behavior as both viewed the videotaped performance; (3) the experimenter both reinforced the desired behavior and pointed out its effects on pupil behavior.

One of the dependent variables was the number of reinforcements given by the teacher each time a student participated verbally in class discussion. The other dependent variable was the number of student participatory responses. Both variables were quantified by counting their number using the videotapes of teacher performances.

It was predicted that providing both reinforcement and discrimination training would produce the greatest change in teacher behavior; and that student behavior would vary directly with the number of reinforcements given by the teacher.

Teaching interns were randomly assigned to these three treatments and to a control group in which subjects rated their own teaching performance on a variety of characteristics. Three experimental sessions intervened between the pre and post-test.

Reinforcement and discrimination training by the experimenter produce the greatest change. Both reinforcement conditions produced significantly more change than the other conditions. An increase of fifteen per cent in number of reinforcements given by the teacher doubled the number of student responses.

Dwight W. Allen, Frederick J. McDonald, and Michael E. J. Orme, Stanford University. The Effects of Feedback and Practice Conditions on the Acquisition of a Teaching Behavior.

Two independent variables were manipulated, kind of feedback and conditions of practice. The two kinds of feedback were immediate, given as soon as possible after a teaching performance, and delayed, given one week after the teaching performance. The amount of practice between feedback sessions varied from none to two weeks.

The dependent variable was a questioning technique which took a variety of specific forms. The purpose of the technique was to elicit extended student thinking about his answers to questions.

Two predictions were made: (1) that immediate feedback would be more effective than delayed feedback; (2) that distributed practice would be more effective than massed practice.

Teaching interns were randomly assigned to four conditions in a two by two design. Three experimental sessions and a pre and post-test were videotaped. During experimental sessions, an experimenter viewed the tape with the intern. He reinforced instances of the desired behavior, pointed out salient cues to which the desired behavior should be attached, made suggestions about variations in the form of the desired behavior. The dependent variable was quantified by counting instances of it from the videotapes.

Immediate feedback was found to be significantly more effective than delayed feedback. However, the conditions of practice interacted with the kind of feedback; for example, immediate feedback and massed practice was more effective than other combinations.

Michael E. J. Orme, Frederick J. McDonald, and Dwight W. Allen, Stanford University. Effect of Modeling and Feedback Variables on the Acquisition of a Complex Teaching Skill.

Two variables were manipulated in this experiment: (1) demonstration of a complex teaching behavior using teacher models; (2) varying the kind of feedback provided on a teaching performance. Four experimental groups were exposed to teachers acting as models demonstrating the desired behavior; two other groups were not exposed to these models. Feedback was varied by having subjects either evaluate their own performance or by having an experimenter view the subject's performance with him, reinforce the desired responses, point out cues to which they should be attached, and suggest variations in the form of the desired behavior. Six experimental groups were created in a 2 x 3 design. Teaching interns were randomly assigned to these conditions.

The dependent variable was a probing behavior in which the teacher requires students to elaborate superficial, first responses to teacher questions and comments. Instances of the dependent variable were counted from the videotaped performances of the subject.

One hypothesis was that rate and level of learning would be significantly higher in modeling conditions. A second hypothesis was that prompting feedback would be more effective than confirming feedback, and that a combination of the two kinds would be more effective than either alone.

A pretest, posttest, and two experimental sessions were conducted. The teaching sessions were in small groups of four students, the lesson lasting five minutes (sufficient length for an adequate number of the desired responses to occur). Teacher models had been videotaped teaching under the same conditions and after training.

Both hypotheses were confirmed. Interaction effects were obtained; for example, a combination of viewing a model with prompting and confirming feedback from the experimenter was the most effective condition.

SESSION 61

Paper session --THE ENVIRONMENT FOR LEARNING

Charles M. Peccolo, Kansas State University. Thermal Environment and Learning.

The study was carried out in a modern two-room research school where thermal conditions were independently controlled and measured in each room. Each room completely equipped to allow normal teaching program to be carried out as part of a research program to determine the effects of various thermal conditions on student occupants.

Pairs of fourth grade children of a suburban consolidated school were matched on basis of achievement, intelligence, age, height, weight, and occupation of parents. Subjects had no knowledge of thermal conditions in room occupied. Half hour in morning and afternoon was required to perform experimental tasks with remainder of the day spent in normal daily elementary school program.

Data seem to justify these conclusions: 1) Significantly higher gains made by pupils in experimental group indicated that prescribed optimum thermal classroom environment was superior to regular thermal environment for all reasoning and some clerical tasks; 2) Interaction between trials and levels and treatments indicated that the prescribed optimum thermal classroom environment favored experimental group in all tasks, although experimental effect varied in some tasks from level to level.

George I. Brown, University of California, Santa Barbara. Operational Creativity: A Strategy for Teacher Change.

The project attempted to bring about significant improvement in teaching quality through a focus on ways to develop creative thinking in the classroom. Three change agent models essentials for expediting the strategy are described including the within-school, the district-wide, and the external change agent. Following the strategy, in order for creativity to become a genuine educational goal for teachers a series of experiences were structured that 1) helped teachers confront aspects of the creative process at both intellectual and emotional levels, 2) provided the psychological security necessary for the teacher to alter her professional life, 3) opened avenues for professional communication and 4) supported innovation in individual and staff operations which includes experimentation with a more open-systemmed orientation to teaching and an increased emphasis on divergent thinking as a curriculum goal.

Attitude tests, autobiographical reports, interviews and measures of preference for complexity, The Welsh Figure Preference Test and The Barron Complexity Scale were used to evaluate teacher change.

The relative effectiveness of the project strategy model and the change agent roles are examined in terms of the possibility of adaption for use by administrators who wish to institute teacher change.

SESSION 61

James B. Kenney, University of Georgia, Athens. A Comparison of the Organizational Climates of Negro and White Elementary Schools with Concomitant Implications for School Administrators.

This study sought to 1) determine if faculties of Negro and white elementary schools in a predominately segregated school system perceived the organizational climate of their schools differently and 2) if they did have different perceptions of their schools, wherein did the difference lie.

The Organizational Climate Description Questionnaire developed by Halpin and Croft was used as the data gathering instrument. The sample consisted of 111 elementary schools in an urban school district in a Southeastern state. Forty-five of the sample schools were Negro and 66 were white.

Responses from 2321 teachers and principals in the study were factor analyzed and yielded a three-factor pattern almost identical to the patterns obtained by the originators of the instrument. Thus, a sociologically homogeneous sample yielded virtually the same factor patterns as did a sociologically heterogeneous population in the original factor analysis. Negro and white faculties differed significantly (.01 level) in their perception of the organizational climate of their schools. A large majority of Negro faculties perceived their schools as either Closed or Paternalistic (as defined by the model presented in the paper proper), while white faculties saw their schools as Open or Paternalistic.

Both white and Negro faculties saw the Organizational Climate of their school as Paternalistic in the majority of cases.

Norman J. Boyan, Stanford University. The Relationship of Extrinsic and Intrinsic Job Satisfaction to Teacher Performance.

The purpose of the study was to explore the relationship of teacher satisfaction and performance. The independent variables were extrinsic and intrinsic satisfaction. The dependent variable was teacher performance. Subjects of the study were 100+ teacher-interns in the Stanford University Secondary Teacher Education Program.

The measure of extrinsic satisfaction was the score on a scale designed to elicit reactions to environmental factors related to teaching. The measure of intrinsic satisfaction was the score on a scale designed to elicit reactions to elements of self-actualization in teaching. Measures of performance were secured from three sources: university supervisors, resident supervisors, and pupils. The instrument for performance measures was the Stanford Teacher Competence Appraisal Guide, which assesses teacher performance in 13 categories.

The data were analyzed for relationships through the use of multiple and partial correlations. Each of the 13 performance categories was used as a separate dependent variable. Results: 1) Multiple correlation of extrinsic and intrinsic satisfaction with performance: a) resident supervisors as raters - statistically significant correlations (.05 level), ranging from .25 to .37, in 11 performance categories; b) university supervisors as raters - statistically significant correlations (.05 level) in 6 performance categories; c) pupils as raters - no statistically significant correlations; 2) Additional results reported for partial correlations.

SESSION 62

Paper session --THE VALIDATION OF EDUCATIONAL MEASURES

J. Douglas Ayers, University of Alberta. *Justification of Bloom's Taxonomy by Factor Analysis.*

The fact that there have been no reported factor analytic studies of items on achievement tests is surprising especially in view of the fact that it would be highly desirable to have homogeneous sub-scores based on content or on objectives.

In this study, the factor analyses were conducted on a 40-item, 4-option, multiple choice test administered to 297 students immediately following completion of a Grade 10 Science program on the Gas Laws. As the students were taught by a common programmed instruction unit, an a priori classification of the items according to Bloom's Taxonomy was made to provide a criteria for rotation of the factors extracted and to check on the adequacy of the a priori classification.

Phi and tetrachoric interitem correlations were computed for 40 items. Two easy items were subsequently dropped because they interfered with the adequacy of the factorization. Unities were placed in the diagonal of the correlation practices and the principal components method used to extract 15 and 16 factors.

Robert M. Pruzek, State University of New York at Albany. *Relationship Between Categorical Factor Analysis and Conventional Factor Analysis with Application to a Study of Two CEEB Achievement Tests in Mathematics.*

Categorical factor analysis (CatFA) is examined as a method for studying the content of mental tests. Given several categorizations of a set of test items, CatFA may be used to study relationships among the different categorizations. The major structural hypothesis for CatFA is that latent categories underlie manifest categories; it is assumed that each test item belongs to a latent category and that manifest categories are derived by dividing and combining latent categories. When the hypothesis is found to be valid, a clear operational definition of the term "parallel tests" is possible. Implications of such analysis for test construction is discussed.

Relationships between CatFA and the conventional principal component and common factor models of factor analysis are examined both algebraically and empirically. Ten CEEB subject-matter specialists in mathematics have independently sorted items from two forms of the College Board's Achievement Series in Mathematics into disjoint categories on the basis of the perceived likelihood of items' measuring the "same" ability or knowledge. CatFA is used for analysis. For response data, two 1400 case samples of College Board candidates are used - one for each test form. The data are analyzed using both the principal component and common factor models and oblique analytical rotations. Relationships between results of CatFA and the two conventional factor analyses are discussed.

SESSION 62

Manford J. Ferris, University of California, Berkeley. *Validity as a Function of Empirical Scaling of Test Items by a Logistic Model.*

The research hypothesis was that correlations between predictor and criterion scores based on empirically weighted items are higher than correlations between test scores each expressed merely as the number of items correctly solved. The quantitative section of the Cooperative School and College Ability Tests (SCAT) Form 2A was used as the predictor measure; the mathematics section of the Cooperative Sequential Tests of Educational Progress (STEP) Form 2A was the criterion test. Subjects were 1,893 eleventh grade students who had completed all items on the two standardized tests. The index of difficulty or limen of a test item was defined as the total test (raw) score for which the probability of a student's having passed the item is 50 percent, where total test score of an examinee is the number of test items he answered correctly. The logistic model showed that five items of the SCAT-Q test had indices of a difficulty less than the minimum expected raw score on the basis of guessing alone; two items of the STEP-M test had low limens and one had a limen larger than the maximum possible raw score (50). In this study, rescoring tests by using item weights computed by the logistic model did not improve predictive validity of the predictor measure. Since there was such a small difference (.055) between the correlation of unit weighted scores on the two tests and the first canonical correlation, use of the logistic model with these tests may not have given the model a fair chance to improve the validity. There may be merit in using the logistic model in the item analysis phase of test construction.

John R. Bormuth, University of California, Los Angeles. *Factor Validity of Close Tests as Measures of Reading Comprehension.*

A 50 item cloze test was made over each of nine passages by replacing every fifth word in each passage with an underlined blank of a standard length. The passages dealt with a variety of subject matters and each was about 255 words in length. Seven multiple choice tests, each designed to measure what were presumed to be different reading comprehension skills, were constructed over the same passages. Construct and statistical validity studies were made in preparing the multiple choice tests. These tests were then administered to 150 subjects enrolled in grades 4, 5, and 6 of an elementary school. The cloze tests were administered first, before the subjects had read the unmutilated versions of the passages. Three days later the subjects were given the multiple choice tests after having been allowed to read the passages. The tests were not timed and subjects were not permitted to refer back to the passages while taking the multiple choice tests. An orthogonal solution of a principal components analysis was calculated. Only one factor emerged having an eigen value greater than one. This factor accounted for 77 percent of the variance. With one exception, the correlations of the test scores with this factor approached the maximum possible for tests of their respective reliabilities. The exception was a test of ability to comprehend the main idea. An examination of the response patterned on this test showed that some subjects had a consistent tendency to select either overly general or overly restricted responses.

SESSION 63

Paper session --THE SCHOOL AND THE LARGER SOCIETY

Jerome Beker, Youth Development Center, Syracuse University. The Evolution of Public Educational Policy: School Desegregation in a Northern City.

This paper reports a study of the attempt made by a medium-sized northern city to deal with the problem of de facto segregation in its public schools. More than a "case study", the report attempts to analyze as well as describe the development of the situation. Data were gathered from minutes of public and private meetings, letters and other documents, local newspaper coverage, confidential depth interviews of over forty key participants in the critical events representing all shades of opinion and influence, and limited participant observation of meetings, public hearing, and the like.

The paper describes the roles played by the primary "actors" and groups involved - such as the Board of Education, the School Administration, the City Administration, the Civil Rights Groups, and the Press - and how the handling of the problem agrees with or deviates from formal conceptualizations of how public educational policy is made. An attempt is made to extract implications and guidelines for other northern communities where similar processes are in earlier stages as well as generalizations about public educational policy formation.

Robert E. Herriott, Florida State University. The Current Population Survey of the U. S. Bureau of the Census as a Vehicle for Educational Research.

Each month, the United State Bureau of the Census conducts a nationwide sample survey called the Current Population Survey (CPS). The primary objectives of the CPS are to obtain a measure of employment and unemployment and to obtain data on the characteristics of the labor force. However, in several particular months, additional data are obtained. Each October, for example, questions are included by the Census Bureau to procure data on the school enrollment of children and young adults in the 35,000 households sampled.

By special arrangement with the Census Bureau, the CPS can be augmented to consider more general questions of national importance. For example, in October 1965 parents will be asked questions about their attitudes toward education and about the educational performance and plans of their school-age children; the older children will be asked questions about their attitudes toward education, and about their plans for further schooling; and the principal of the school attended by each child will be asked questions about the child's performance, as well as about the social and economic composition of the school's student body and staff. These data, when combined with those regularly obtained by the CPS, are expected to provide a valuable source of knowledge about social and economic factors affecting educational opportunity.

SESSION 63

Robert B. Brumbaugh, University of Oregon. Reference Group Orientation and Teacher-Attitude Toward School District Reorganization.

This study investigates resistance to organizational change within the conceptual framework suggested by reference group theory. Specifically, the concept of local-cosmopolitan orientation is applied to 329 public school teachers in five contiguous school districts in Pennsylvania. All districts were faced with imminent mandatory reorganization.

The rationale from which the guiding hypothesis was derived concerns past research findings indicating that locals are generally oriented toward more geographically and psychologically immediate reference groups in contrast to cosmopolitans who characteristically derive their norms from more distant and diffuse sources. In view of this, it was hypothesized that teachers who were locals would be more resistant to school district reorganization than would teachers who were cosmopolitans. Using the Sutthoff local-cosmopolitan index, the study identified 51 teachers as locals and 218 teachers as cosmopolitans, but found no significant differences in the responses of these two groups to a scale measuring attitude toward school district reorganization.

One provocative finding, among others, was that a significantly greater number of those teachers who scored low (less favorable) on the reorganization attitude scale also scored low on a test of factual knowledge concerning the details of the reorganization legislative mandate.

A further finding with direct implications for school administrators was that locals and cosmopolitans characteristically obtain information from significantly different sources. The two-step flow communication process was observed to be sharply in evidence for locals.

Donald A. Erickson, University of Chicago. Social Functions of Nonpublic Schools: The Research Evidence.

Serious conflicts between supporters of public and nonpublic schools occur, but may symptomize deeper factors. Nonpublic schools often display discrete subcultures not necessarily products of the schools themselves. Nonpublic schools probably do not alienate individuals from the societal mainstream. But some minorities may need nonpublic schools to survive.

Important innovations have come from a few nonpublic schools, but most are conservative, even reactionary. Nonpublic schools may have played a more responsive role than public schools in rehabilitating some neighborhoods. Often nonpublic schools have tended to create a pauper school image of public schools. Some nonpublic schools appear to exist to avoid racial and socioeconomic heterogeneity.

Public schools have been deprived of potential resources because nonpublic schools exist. But nonpublic schools have used non-tax resources to educate many students. The net effect seems to be to raise per-pupil expenditures in public schools. Shared time evidence suggests educational programs have been curtailed through the noncooperative maintenance of public and nonpublic schools, but the latter serve certain needs not otherwise met.

On the basis of the evidence, one may rationalize a desirable kind of coexistence and cooperation between public and nonpublic schools, deriving guidelines for policy.

SESSION 64

Paper session --ORGANIZATIONAL VARIABLES: INNOVATION AND EVALUATION

James A. Reynolds, Hoechs Jr. High School, St. Louis County, Missouri. *Innovation Related to Administrative Tenure, Succession and Orientation.*

This study was based upon the assumption that the superintendent is significant in determining the innovation that takes places in the school district where he is employed. The relationship between innovation in elementary and secondary schools and the tenure, succession pattern, and reference group orientation of the superintendent was investigated.

Data pertaining to tenure, succession pattern, and local-cosmopolitan reference group orientation were obtained from questionnaires sent to 183 superintendents in Missouri and Illinois. Information about innovation was gathered from questionnaires sent to elementary and secondary principals in these districts. The principals were asked to provide information about the adoption status of 19 elementary and 23 secondary practices. The final analysis was based upon the complete returns of 94 districts.

Factor analysis was used to develop the scales for local-cosmopolitan reference group orientation, perceived need for change, and perceived power to innovate. The hypotheses were tested by use of a 2 x 2 x 2 analysis of covariance design with expenditure per pupil and size of the district serving as control variables.

The study supported the assumption that the superintendent is significant in determining the adoption of new practices. Innovation is, however, a complex process and there are numerous factors which operate to limit his range of control. The importance and complexity of the innovation process argues for continuing research in this area.

John W. Childs, Wayne State University. *A Study of the Belief Systems of Administrators and Teachers in Innovative and Non-Innovative School Districts*

The purpose of the study was to collect and examine empirical evidence relevant to the idea that there is a relationship between the nature of the belief systems of individuals in school districts and the adoption of new educational practices. Eight school districts were selected for study using a constructed innovativeness scale based on the Michigan Department of Public Instruction's Five Year Survey of Progress in Michigan School Districts. The selection controlled for four cost factors: size, expenditure, state equalized evaluation, and operational millage. Districts were identified as being innovative or non-innovative and as being high cost factor districts or low cost factor districts. The Dogmatism Scale was administered to 755 teachers and 51 administrators. It was hypothesized that the proportion of administrators and teachers having open belief systems in innovative school districts would be greater than the proportion of administrators and teachers having open belief systems in non-innovative school districts. The analysis for the teachers yielded a chi-square of 8.42 which was significant at the .01 level. The research hypothesis suggesting that a greater proportion of teachers in innovative schools would be open minded was supported.

SESSION 64

Herman A. Wallin, University of Oregon, Eugene, Oregon. Incentive Systems in Educational Organizations: *A Look at Two Types and Their Apparent Impact Upon Skill Maintenance Activities of the Professional Work-Force*.

The study examined the impact which different organizational incentive systems had upon the attitudes and activities of the professional staffs of two institutions of higher learning. One of the public junior colleges had a merit-type incentive system; the other did not differentiate among the various contributions of its faculty: the extent of academic preparation and mount of previous teaching experience determined the size of the compensation.

Variations between the two faculty groups in the matter of skill maintenance activities are expanded on the present paper. Three components of this activity are discussed: subscribing to and reading relevant journals, engaging in formal studies, and participating in conferences and seminars.

The college employing the merit provisions appeared to elicit from its professional staff members a greater commitment to maintaining their expertise. A significantly greater proportion of the teachers in the non-merit college discontinued early in the careers any contact with journals or their colleagues at conferences and seminars. Formal studies were often viewed as the means to enhance salary; there was, in many cases, little regard for the relevance of those studies to one's teaching field.

Jacquetta H. Burnett, University of Illinois. *Student Workers and Quality Control in Person-to-Person Work Flow of a Rural High School*.

To clarify the technology of school organization during a participant observation study of a Midwestern high school, the writer borrowed from industrial relations research the concept of a person-to-person work flow. Work flows can be identified wherever there is a sequence of techniques that must be performed in a regular or predetermined order by separate individuals. Two charts characterize the analysis of work processing in the high school in terms of person-to-person action, spatial movement, and temporal periodicity.

After noting that other types of quality control than the statistical type may be used, quality control check points in the school's work flow are discussed. The search for "reprocessing" procedures which bring poor work up to standard shows that the school used highly inefficient procedures where it used any at all.

Observations of student classroom activity and analysis of their responses to questions concerning work and rewards in the academic system suggest that students acted like and viewed themselves as workers. Certain structural features of the organization of work, it is argued, discouraged the teachers from accepting the ideology that the students are products of work or are clients of teachers services. Observation of teachers' actual behavior and verbalizations offer corroboration for that interpretation. The implications of viewing students as workers for present procedures of quality control and and of "reprocessing", and for arrangements for productivity are discussed.

SESSION 65

Paper session - STUDENT ATTITUDES AND ACHIEVEMENT

Milly Cowles, Rutgers - State University of New Jersey, Kathryn B. Daniel, Newark State College, and Corrina Kay, Newark Public Schools (New Jersey).

The problem was to study the effects on I.Q. scores of a plan designed to bridge the cultural and educational gap for disadvantaged junior high school pupils. Comparisons were made between high-low and low-low socio-economic pupils with respect to scores on the Lorge-Thorndike Intelligence Test.

The sample was composed of 192 junior high school pupils. A high percentage of sub-standard dwellings qualify the area as a slum section. A program of expanded opportunities for two years was given the subjects. The program included class trips, cultural programs, additional instructional materials, guidance services, and additional staff members.

The Lorge-Thorndike Intelligence Test was administered at the beginning of 7th grade and at the end of 8th grade. The interval between tests was two academic years. The Lorge-Thorndike scores of the two groups were tested for differences after the first testing. A negligible F-ratio was obtained and it was assumed that I.Q. scores were similar for both socio-economic groups at the beginning of the enrichment period. Differences in pre and post tests for both groups were tested for significance by means of analysis of variance. The F-ratio of .067 did not approach statistical significance. After a period of two years of cultural and academic enrichment, there was no significant difference between the two groups" I.Q. score gains.

Melvin Arnoff, Kent State University. The Attainment of Social Studies Concepts: The Six Year Old.

The concensus of learned opinion is that the social studies curriculum of the public school is no longer an effective instrument for the realization of currently appropriate educational goals. One major complaint is that the fact load of the social studies underestimates children's abilities, especially in the primary grades. It is also argued that children should not be insulated from studies of the world until the 4th year of school. Further curriculum development appears dependent upon assessment of the abilities of 6-year-olds to: 1) develop concepts beyond those of grade one social studies texts, and 2) learn selected concepts about a foreign culture.

Teachers of first grade were identified as "social studies emphasizing", (SSE) and "non-social studies-oriented", (NSS). Three classrooms were selected from each group. A "Practice-Test" ("Families"-Form A) was administered in December. These results were not analyzed. Form B was analyzed in January, 1966. Form I of "Families in Japan" was administered in January, 1966.

These procedures identified: 1) Selected concepts of "Home and Family" learned by children with NSS and SSE teachers, and 2) Selected concepts of "Families in Japan" known prior to instruction. The latter test will serve as a basis for assessing children's abilities to assimilate concepts of a foreign culture and for evaluating an experimental unit on Japan for grade one.

Sally Anthony and Harry V. Barnard, San Diego State College. Contradictions in Beliefs Among Ninth and Twelfth Grade Students.

The first purpose of this exploratory study was to devise an instrument to measure self-contraditory beliefs in the following controversial areas: 1) sex, courtships and marriage; 2) economics; 3) prejudice; 4) nationalism and patriotism; 5) social class; and 6) religion and morality. The second purpose was to investigate differences between 9th and 12th graders and between boys and girls with respect to those contradictions.

The Contradiction Inventory was developed by taking statements from the literature which were said by psychologists, sociologists, and educationists to represent common American beliefs. Sixty-six pairs of contradictory statements were compiled. Five judges passed unanimous agreement on the contradictory nature of 52 of the pairs. These 52 pairs were separated into 104 statements. Five new judges were asked to fit the statements together into contradictory pairs. Eighty-six percent agreement was received. Ten judges were then asked to determine whether the statements fitted into one of the 6 controversial subject areas. Fifty-one pairs of statements received at least 90% agreement. These 51 pairs of statements make up the final form of the Contradiction Inventory. The coefficient of stability was r=.75, sufficient for purposes of studying groups.

The Contradiction Inventory was administered to 118 students. There was no significant difference in self-contradiction between 9th and 12th grade students and no significant difference between boys and girls.

Joseph Jackson, James Creel, and Harold Myler, Dearborn Public Schools, Michigan. Attitude and Skills Changes as a Result of Humanities Emphasis in Communication.

While academic gains are identified through achievement tests, awareness of growth in the attitudes and feelings of students is subjectively appraised. Under trial in this study was the oft-repeated approach that the study of a country involves a conglomeration of facts about hills, mountains, boundaries, exports, etc. To provide maximum appreciation of the aesthetic and of the spiritual, empathic indulgence was provided through music, dances, food, dress, crafts and art. The study began with a unit on Italy at grade 6. Evaluation and growth appraisals were made possible through pre-and-post themes, the basic title being "What is a Country". Growth in communications skills for the group was had through the local, "Marking Standards for Written Communication".

The pre-test theme approached the task through historic concepts. The retest set of papers retained the obvious ideas but did add comments of human values and of contributions. Attitudes do not appear to have a planned continuity beyond the teacher of interest. A directed emphasis upon humanistic concepts must begin before grade 6.

SESSION 66

Paper session --SPEECH AND LANGUAGE LEARNING IN CHILDREN

William J. Griffin, George Peabody College for Teachers. A Transformational Analysis of the Language of Kindergarten and Elementary School Children.

Comparable language samples were collected on tape from 180 boys and girls about equally distributed in kindergarten and Grades 1, 2, 3, 5, and 7; written samples were also elicited from Ss in grades 3, 5, and 7. Mean IQ's and achievement scores within grades were closely comparable. The corpus was submitted to various measurements; most important were analyses of syntactic features produced by sentence combining transformations.

Sensitive general indices of development toward linguistic maturity were in both the mean length of single independent predications and the mean number of sentence-combining transformations contained in them. Particular syntactic features which best reflected language control were transformations involving deletion rules. Greatest gains in syntactic variety and complexity in speech were evidenced in Grades 1 and 7; speech development in the middle grades seemed relatively slight. In writing, most dramatic growth was in Grades 4 and 5 though Grade 7 showed substantial advances. Syntactic control in writing, much inferior to speech in Grade 3, improved rapidly so that in Grade 7 it was superior to speech.

Girls were superior to boys in written expression in Grades 3 and 5, but not in Grade 7. In speech, linguistic development of boys appeared in several significant ways to outrun that of girls in the same grades.

Melvyn I. Semmel and Beverly G. Herzog, University of Michigan. Verbal Recall of Negro and White Educable Retarded Children as a Function of Grammatical Form Class.

The purpose was to study the effects of grammatical form class on the verbal recall of Negro and White Educable Retarded children. Ss were ten Negro and ten White boys aged 9-1 to 12-6 controlled for IQ and socio-economic background. Each S was given a verbal recall task composed of 20 words which were of the same frequency of occurrence among retarded and normal children of the same MA. (Mein Word List, 1961). Five nouns, verbs, adjectives, and adverbs were randomly presented via tape recording over three trials. The words recalled by the S were recorded. The S was asked to give a sentence using those words (adjectives and adverbs) which might belong to more than one form class. Form class was found in this study to be a significant variable in verbal recall. Nouns were recalled more often and adverbs least often. There was no significant differences between adjectives and verbs. Negro Ss had a significantly higher recall than White Ss. Ss recalled significantly more words on the 2nd and 3rd trials than the 1st. Interaction effects among the variables were not significant.

SESSION 66

Melvyn I. Semmel, L.S. Barritt, S.W. Bennett, and C.A. Perfetti, Univ. of Michigan. *A Syntactic Analysis of Retardate Free-Associate Responding.*

The purpose was to examine differences in paradigmatic free-associate responding by normal and retarded subjects. The sample consisted of 20 Institutionalized Retardates (IR), 20 Public School Retardates (PSR), 20 normal subjects matched with the retarded groups on Mental Age (MA), and 20 normal subjects matched with the retarded groups on Chronological Age (CA). Forty words were selected from the Mein and O'Connor (1961) list. There were 10 nouns, 10 verbs, 6 adjectives, 5 pronouns, 5 adverbs, and 4 prepositions. An Analysis of variance showed that paradigmatic (non-sequential-Homogeneous) responses occurred more frequently in the normal CA match group than in any other of the sub-groups. The PSR and MA matched groups produced the next greatest number of paradigmatic responses while the institutionalized retarded group gave the fewest. Across all form classes, the MA group was essentially equivalent to the PSR group in the number of paradigmatic responses. Normal and retarded Ss showed the greatest similarity in responses to nouns, although retardates and MA matched Ss still produced fewer paradigmatic responses than the CA normals. The results suggest that with increasing chronological and mental age, children move from a predominately sequential approach to the relationships between words, to a "higher" level of language development in which associations between words having similar privileges of occurrence takes on equal importance.

Donald E. Mower, Robert L. Baker, and Richard E. Schutz, Arizona State University. *Of Meith, Myth, and Men; Extending the Control of Lisping Behavior to Non-Laboratory Stimuli.*

Previous research in our Laboratory has developed programmed learning procedures for the elimination of lisping behavior. However, this control over lisping did not extend to free conversation. The present study investigated procedures for extending the stimulus control to non-laboratory environmental stimuli. A 2x2 design was used, with the person monitoring the behavior and the reinforcement contingencies as the two main effects. The child's mother was used as the monitor under one condition while both mother and the classroom teacher served as monitor under the other condition. A redeemable token or no token represented the reinforcement. Eighteen kindergarten children diagnosed as "lispers" completed a three-session laboratory training program, designed to elicit and maintain correct S-responses in the presence of the visual and auditory stimuli presented in the training sessions. The children were then randomly assigned to one of the 4 experimental treatment groups. The teacher and/or parent administered a 3-week speech correction program designed to maintain behavior and extend stimulus control. "Conversational" responses cued by picture cards depicting children's activities were used as the criterion measure. Results were: 1) children reinforced for correct responses produced sig. more correct Ss than those not reinforced; 2) the parent-teacher and the parent-only conditions did not differ sig. There was no interaction between the two effects.

SESSION 67

Paper session -- ASSESSMENT OF EDUCATIONAL ENVIRONMENTS

Alan M. Abrams, Stanford University and University of Wisconsin, and Julian C. Stanley, Center for Advanced Study in the Behavioral Sciences and University of Wisconsin. Certain Relationships Between Characteristics of Students and Their Colleges.

Data for more than 1000 colleges collected by Alexander W. Astin were analyzed to see what relationships exist between eight characteristics of the colleges and five characteristics of their beginning freshmen. The variables defining "characteristics of students" were measured or recorded as the student entered the college; on the other hand, college characteristics were generally defined by aggregate measures of the previous outgoing student bodies.

The technique of canonical correlation is employed to measure the relationship between these two sets of variables. This measure is compared with other measures of relationship, and it is argued that canonical correlation constitutes a meaningful, supplemental measure of relationship in this situation.

Of special interest are the differences among these relationships when they are studied separately among different types of students and colleges. That is, do these relationships differ between Negro and white colleges?

Suggestions are made regarding how this type of study should be done in order to make the kind of inferences which one wants to make, since causal inferences cannot be made directly from correlational studies. The question of how one might proceed to answer the question, "What effect does a college have on a person who attends that college?" is discussed. Similarly, the question of how one measures the effect of a person on a college is considered.

Basically, the paper is a summary of the findings resulting from analyses of data and a careful interpretation of the findings.

Edwin B. Hutchins, Association of American Medical Colleges and Arthur J. Nonneman, University of Michigan. Construct Validity of an Environmental Assessment Technique for Medical Schools.

Interest in interaction process variables includes the work on student environments by Pace and Stern; Nunally, Thistlethwaite and Wolfe; Astin; and Halpin and Crofts. The development of the Medical School Environment Inventory (MSEI) by Hutchins represents an extension of this effort to the medical college setting. Developed in the context of a longitudinal study of 28 medical school classes entering in 1956, the MSEI has frequently been used in applied setting though often uncritically. The purpose of this paper is to present follow-up data bearing on the psychometric characteristics of the MSEI and on its utility as a research and applied educational tool.

Estimates of the internal consistency of the MSEI scales obtained from the original sample ranged from .88 to .99. The test-retest reliabilities over a five-year period which range from .76 to .98 are presented along with faculty-student and class-to-class comparisons. The construct validity of the test is explored by presenting a nomological network of relationships with other variables such as the value structure of the student body, measured achievement motivation, and premedical school environment data obtained from Astin. High concurrent validities were obtained with career choice and attrition information in 1960, and these are compared with predictive validities obtained on the same group five years later after the majority had completed their residency training.

James V. Mitchell, Jr., The University of Rochester. The Identification of Student Personality Characteristics Related to Perceptions of the School Environment.

Current research on the characteristics of school learning environments raises the important issue of whether student perceptions of the school environment are independent of personality characteristics that might bias these perceptions. The purpose of the present study is to determine whether such relationships do exist, to define their magnitude and nature, and to assess their significance.

Administered to the 223 members of the junior class of a suburban high school were the High School Characteristics Index, the California Psychological Inventory, the SRA Youth Inventory, and the Survey of Study Habits and Attitudes. Correlations were then computed between the HSCI and the CPI, SRA, and SSHA. The results indicated that:

1) There are significant and appreciable relationships between personality traits and environmental perceptions, with perceptions of some aspects of the school environment (e.g. HSCI Abasement and Objectivity) exhibiting stronger relationships than others.

2) The CPI Achievement via Conformance score was the personality variable most highly related to perceptions of the school environment. Its highest correlations included a negative relationship with the number of problems checked on the SRA "My School" scale and additional negative relationships with HSCI scales representing environmental Abasement, Aggression, and lack of Objectivity.

3) Of the SRA scales, the "Home and Family" scale had the highest correlations with environmental perceptions. The total number of problems checked in this area was most highly related to HSCI scores representing environmental Abasement and lack of Objectivity.

The results are interpreted as suggesting that perceptions of the school environment are particularly influenced by student conformance to school achievement values or rebellion thereto, that non-conformance is related to processes serving to distort the student's perception of the intent and fairness of persons in the school environment, and that experiences in the home may set the pattern for the student's perception of all his environmental experiences, including those in the school.

SESSION 68

Paper session --INTELLECTUAL DEVELOPMENT AND ITS MEASUREMENT

Anne M. Bussis and Richard S. Melton, Educational Testing Service. Piaget for First Grade Teachers: A Guide for Understanding Intellectual Development.

A guide for teachers was prepared indicating items significant for cognitive development of first grade children. Items chosen were those which would indicate development to the Piagetian level of logical concrete functioning. These items were organized into six areas: concepts of space and time, growth of logical reasoning, understanding mathematics, oral communication, learning about the world, and imagination and creativity. After the first year's try-out in a sample of 25 New York City schools, the items were reorganized into the following six areas: basic language skills, concepts of space and time, beginning logical concepts, beginning mathematical concepts, the growth of reasoning skills, and general signs of development. In each area, suggestions for instructional and assessment tasks are given to the teacher, with references to a companion manual: Instructional and Assessment Tasks, as well as to a manual which provides further elaboration of the theory on which manuals are based. From Theory to the Classroom: Background Information for the New York City Project.

Masako Tanaka, Joel T. Campbell and John S. Helmick, Educational Testing Service. Piaget for First Grade Teachers: Written Exercises for Assessing Intellectual Development.

Six sets of written exercises were prepared for use in first grade classes, as part of a larger project undertaken for the New York City Board of Education. The primary purpose of the exercises was to provide an assessment of the entering pupil's ability early in the school year. Thus, provision had to be made for limited knowledge of pencil and paper work. A solution was to provide in each area three days of practice work and two days of measurement.

Exercises have been constructed in six areas: shapes and forms, spatial relations, time concepts, communication skills, logical reasoning and mathematical understanding. Piagetian principles were used where they could be applied, and items chosen from those likely to be in the urban child's repertoire. Preliminary results indicate that the practice sessions are probably teaching concepts as well as test-taking skills.

Elliot W. Eisner, Stanford University. The Development and Use of a Scale for Assessing Space in Children's Drawings.

The purpose of this study was to construct a scale that would be useful for assessing the ways in which children treat space in their drawings. To construct this scale approximately 1,500 subjects attending seven schools in and around Chicago produced crayon drawings with standard materials under standardized conditions. The subjects were in grades 1, 3, 5, and 7, and were divided equally among Negroes and whites. One-half of each racial group came from slum communities and was considered culturally-disadvantaged, and one-half came from middle-class communities in cities and suburbs and was considered culturally-advantaged. From previous research in the field of child art, four general stages describing the ways children draw space were identified. Analyses of the collected drawings made it possible to construct ten additional categories bringing the number of categories which constitute the scale to fourteen. Once these categories were selected, verbal criteria were formulated for each of them, and visual exemplars were selected.

To test the objectivity of the scale, two judges were asked to independently sort the 1,500 drawings into the categories they deemed appropriate as represented by the scale. It was found that the judges agreed on 71 per cent of the drawings. The drawings on which the judges disagreed were identified and the judges were then asked to view these drawings again, and to try to come to agreement if they could. On this joint evaluation procedure, 93 percent of the 29 percent of the drawings on which they previously disagreed were categorized.

The drawings which were categorized either independently or jointly were then subjected to a series of analyses. These analyses were made to identify sex differences, differences in development between Negro and white sub-groups, and to identify differences among levels of development between culturally-advantaged and culturally-disadvantaged children at each grade level.

Aaron S. Carton, Division of Teacher Education, City University of New York

A Visual Inference Test: Exploratory Study

A test to measure both the ability and willingness with which pupils make inferences was developed in a USOE-sponsored research project on the "Method of Inference" in Foreign Language Learning. An inference was defined as a response to an unfamiliar stimulus on the basis of familiar attributes of the stimulus or the context of occurrence.

The test-task required Ss to identify small segments - called "targets"- of pictures which were masked to make the targets generally unrecognizable. Five additional segments of the picture provided successive and cumulative hints about the target. With each attempted identification, occurring upon the presentation of an additional clue, Ss supplied "certainty ratings" on a five-point scale. No instructions about scoring were given with the first group of items. Prior to the second group of items, Ss were told that their certainty ratings would determine how much credit would be added or subtracted for correct or incorrect responses.

Test scores provide data on the number of correct identifications, incorrect attempted identifications, general level of certainty, and the relationship between certainty ratings, correctness of response, and number of clues supplied. The magnitude and variability of certainty ratings under the two scoring conditions reflect certain aspects of risk-taking behavior.

In this exploratory study, validation of the several test scores is in terms of a) teacher-ratings of inference behavior while learning French, b) Cloze tests, c) IQ measures, d) school achievement measures, and e) tests of mental functioning. Data from 3 junior high school samples, ranging from 83 to 232, will be presented.

SESSION 71

Paper session--MEASUREMENT OF CLASSROOM PROCESS VARIABLES

Oren Glick, Community Studies, Inc. The Educational Process in the Classroom: Some Problems in Conceptualization and Basic Assumptions.

This paper focuses on some of the assumptions about the educational process as it takes place in the classroom which underlie a great deal of educational research, teacher training programs, and teacher classroom behavior. It is accepted that: (1) the teacher is somehow important and (2) outcomes in terms of individual pupil attributes are important. The paper is concerned with the problems arising in the very next step in the logical development of a model. How does something having to do with the teacher have effects on something having to do with individual pupils? How might the form of the connecting link between these two entities be most profitably conceptualized?

The assumption typically made is that the causal linkage in question is direct and one to one, that it consists of numerous discreet and independent teacher-pupil connections and that the causal direction, as far as the educational process is concerned, is from teacher to pupil. It is argued that the assumption is inadequate since it fails to find sufficient correspondence in the realities of the classroom situation.

An alternative formulation is provided in which it is assumed that teacher effects are mediated by conditions and processes of the classroom peer group, the latter becoming intervening variables between teacher behavior and pupil outcomes. Schematic representations of the two models are provided.

Anita Simon, Thomas Samph, Robert Soar, and Edmund Amidon, Temple University. Programing Teacher-Pupil Interaction Patterns.

The purpose of this paper is to report on the development and testing of a computer program designed to identify specific elements of teacher-pupil interaction patterns which can be used to differentiate teachers or groups of teachers from one another.

Flanders Interaction Analysis data was collected by trained observers on a sample of 88 teachers. These teachers were placed in four groups each receiving different treatments. Three direct categories (lecture, commands, criticism) and four indirect categories (acceptance of feelings, praise, acceptance of ideas and questions) were used in observing the teachers.

A program developed by Soar was used to convert the raw data into matrix form. This matrix form allows the reader to see not only what behaviors were exhibited by students and teachers, but which behaviors followed which behaviors.

The findings indicate that some gross measures, such as the proportion of direct to indirect behaviors, did not differentiate significantly among groups. Also, a large number of specific measures did not differentiate significantly among groups, indicating consistency among groups of teachers in these behavior patterns. There are, however, certain behavior sequences which do discriminate among groups of teachers. Among the discriminating factors are those which represent sequences of behaviors in which the teacher is immediately responding to a student's comment, the amount of extended student talk, and the extended teacher use of a student's ideas.

This computer program provides a convenient method for processing large amounts of Interaction Analysis data along many dimensions.

Jenny R. Armstrong, and M. Vere DeVault, University of Wisconsin and Eleanore Larson, University of Rochester. Consistency of Teacher Communication: A Sampling Problem in Interaction Analysis.

A major problem confronting researchers and curriculum workers interested in interaction analysis techniques in education is that of determining the time required to adequately sample a teacher's communication behavior. This study was undertaken to provide some tentative answers to this persistent problem.

Four elementary teachers were tape recorded with a vega microphone for an entire week. Data used in the analysis included all teacher communication behavior collected during the week for each of the four teachers in reading, social studies, arithmetic, and language arts. A category system developed in the Wisconsin Teacher Education Research Project and the Flanders system were used in analyzing the recorded data. Chi square analyses between consecutive communication matrices (based on multiples of ten minutes) and the matrix for the total week provide data concerning the time beyond which classroom observation must be continued to obtain an acceptable level of confidence. This acceptable level may vary according to the task for which the observation is made (e.g. in-service education or research). Results, therefore, are reported for various confidence levels.

Results indicate that for both the Wisconsin TERP and Flanders systems a time period of approximately two and a half hours of observation time is necessary if a confidence level of .05 seems acceptable.

John Herbert, Reed College. The Analysis of Lessons.

This report describes a new system for analyzing the teaching of lessons. The analysis of the teaching of lessons requires first that the giving of lessons be distinguished from other ways of teaching, such as teaching by creating an environment, teaching by providing a living example, or teaching by fostering meditation, which require a different system of analysis. A lesson is taught whenever there is an instructional relationship between a student and a teacher able to change the 6 analytically independent components of the lesson which are: subject matter, the form of the subject matter, the form of the lesson, the media of the lesson, the grouping and location of students and teachers, and the influence techniques used. A number of discrete forms in which each component occurs in teaching was identified. Categories of these forms were identified by simple criteria, and lessons or portions of lessons were described by charting changes in the components. Individual teachers and teams of teachers were found to have characteristic patterns of choices of components, not identifiable without this systematic analysis.

SESSION 72

Paper session --THE CULTURALLY DISADVANTAGED STUDENT.

Max Weiner, Brooklyn College , N. Y. C. , and Shirley Feldmann, City College, N. Y. C. *Problems in Construction and Validation of a Test for Children from Different Socio-Economic Levels.*

Reports on research and related problems encountered in the construction and validation of the Reading Prognosis Test. This test provides a diagnostic information on skills underlying reading as well as predictive scores on future reading achievement for both low and middle socio-economic status (SES) populations. Three major problems were encountered.

The first problem is that of matching children from different socio-economic levels and geographical areas. Scales used to categorize children into socio-economic levels did not adequately identify inter-subject differences within cells, that is, low socio-economic children from different geographical areas were not comparable.

The second problem was concerned with the establishment of a sufficiently low base for each subtest in order to insure that all Ss independent of SES and geographical area would attain at least a basal score.

The third problem dealt with the interpretation of achievement on standardized reading tests which were used as criterion measures. A reading grade equivalent earned at the lower end of the score distribution may represent either a chance score or a minimal reading ability. Even scores at the high end of the distribution presents difficulty in interpreting the quality of the Ss' reading ability.

This paper reports on the data collected in four validation and two follow-up studies.

Julian C. Stanley, Center for Advanced Study in the Behavioral Sciences, University of Wisconsin and Angela B. Biaggio and Alan M. Abrams, University of Wisconsin. *Relative Predictability of Freshman Grade-Point Averages from SAT Scores in Negro and White Southern Colleges.*

Frequently it is stated sweepingly that ability tests are "not valid" for testing culturally disadvantaged persons, particularly culturally disadvantaged Negroes. There are at least four main types of validity, one of the most important being predictive validity. Evidence is accumulating that certain scholastic-aptitude tests predict college-freshman grades at least as well for students in predominantly Negro colleges as for students in predominantly non-Negro colleges. In the present study, we report extensive analyses of freshman data for five academic years at the eighteen state colleges of Georgia, three of which still enroll very few non-Negroes. By subjecting John R. Hills's published data to a number of analyses of variance, we conclude that freshman grades in the Negro colleges are predicted correlationally at least as well as are grades in the predominantly non-Negro colleges when the drastically restricted range of scores of Negroes on the College Entrance Examination Board's Scholastic Aptitude Tests (SAT), Verbal and Quantitative, is taken into account, and not significantly differently for Negro and non-Negro *males* even when the restricted range is not. Also, the two races do not differ significantly with respect to variance errors of estimate based on the best-weighted linear composite of SAT-V, SAT-M, and high-school average for predicting freshman-year grades. Most of the differences between colleges within races are significant. The year effect is sometimes significant for women. Race sometimes interacts with year, chiefly for women. Possible reasons for these results are considered, and implications for testing and curriculum change are drawn.

Albert J. Harris, Division of Teacher Education, City University of New York, Robert J. Lovinger, City College , N. Y. C. , and Jacob Cohen, New York University. Factor Analysis of Intellectual Performance in Disadvantaged Negro Adolescents.

The WISC and Metropolitan Achievement Test (Intermediate, Complete) were administered to the entire entering seventh grade of a New York City junior high school in a depressed area with an almost 100% Negro population. Full scale and subtest scores are presented and analyzed. Factor analyses of the WISC, the Metropolitan, and the two scales combined have been carried out, using an oblique rotation. The results will be presented and compared to those obtained by Cohen, using the original WISC standardization population.

J. M. Regal, Department of Human Resources, Oakland, California. An Evaluation of Programs Directed at a Negro Population in a Depressed Community.

In March, 1962, the Ford Foundation made a grant of $2,000,000 to the City of Oakland to assist the Community in undertaking a program of social intervention. The findings from the following demonstration projects will be available for reporting at the AERA conference:

OAKLAND PUBLIC SCHOOLS. 1. Language Enrichment Program at a Child Care Center; 2. Kindergarten Language Awareness Program; 3. First Grade Reading and Language Development Program; 4. Third and Fourth Grade Language Development Program; 5. Special Instructional Programs at a Junior High School; 6. After-School Study Centers; 7. Elementary School Library Program; 8. Elementary and Junior High School Counseling Program; and 9. In-School Special Counseling Services to 11th and 12th Grade Students.
RECREATION DEPARTMENT: An Experiment in Intergroup Activities.
HEALTH DEPARTMENT: Health Maintenance Among Newcomers.
COUNCIL OF SOCIAL PLANNING: District Community Councils.
ASSOCIATED AGENCIES: Group Counseling for Boys in Elementary Schools and Coordination of Agencies.
BAY AREA URBAN LEAGUE: Leadership Development Program.
ALAMEDA COUNTY PROBATION DEPARTMENT: Pre-Trial Release Project.
Teenage Pregnancies Program.
Adult Minority Employment.

The presentation will include a brief description of each program, our findings, and some observations relating input to outcome. If time permits, the role of a research department in a community action program will also be discussed.

SESSION 73

Paper session --MEASUREMENT IN CURRICULUM AND INSTRUCTION

Leopold E. Klopfer, University of Chicago. A Semantic Differential Approach to the Assessment of Student's Images of Science, Scientists, and Science Instruction.

Students' images of concepts related to science and science instructional activities clearly play some role in promoting or inhibiting learning in science. Moreover, the establishment of favorable images of science and scientists is one of the important objectives of science education. The semantic differential developed by Osgood, though hitherto little used by researchers in science education, provides a promising technique for assessing the students' meaning structure of concepts relevant to the teaching of science.

Several semantic differential instruments have been constructed to investigate students' images of various concepts, including "Astronomy", "Biologist", "Chemist", "Doing Science Experiments", and others. Five-point bipolar adjectival scales were used. These instruments were administered to some 1500 children at different grade levels, between grades 5 and 12. Factor analyses of the responses on the bipolar adjectival scales have disclosed rather interesting meaning structures for most of the concepts tested at each grade level. Factors akin to Osgood's Evaluative factor are persistent throughout, but other factors, such as an Enjoyment factor and a Friendliness factor, appear to be closely related to the particular concepts used and the age of the students.

David W. Stickell, The Pennsylvania State University. A Comparison of the Achievement Test Performance of BSCS and Non-BSCS Biology Students.

From data collected on candidates who took the College Board Biology Achievement Test in May 1963, analyses were carried out to test the significance of differences between the adjusted mean score on Biology of non-BSCS, or "conventional", students and each of three other groups of students who had taken one of the courses prepared by the Biological Sciences Curriculum Study: 1) the Blue version, 2) the Green version, and 3) the Yellow version. Analyses were based on students who took the Biology Achievement Test in May 1963; Biology scores were adjusted for differences on SAT-Verbal, on SAT-Math, and on a subset of Biology questions that were judged to be appropriate for both the students who took a "conventional" biology course and those who took a particular BSCS course.

After adjustment, there was a difference of 18 points on the College Board scale in favor of the Yellow version students over "conventional" students. There were differences of 12 and 10 points in favor of the "conventional" students over the Blue and Green version students, respectively. Only the difference in favor of the Yellow version students was significant at the .05 level.

There is a discussion of some of the problems encountered in conducting the study and of the relevance of these results to the question of whether more than one College Board Biology Achievement Test should be offered.

SESSION 73

David A. Payne, and Harold A. Vaughn, Syracuse University. Forecasting Italian Language Proficiency of Culturally Immersed Students.

The present study developed as an attempt to identify those variables or combinations which would allow for optimal selection of individuals most likely to benefit from intense language training when immersed in the culture of the language they were studying. The subjects were students, primarily Sophomores and Juniors, in the Syracuse Semester in Italy program (N = 86).

Students received ten clock hours of language instruction each week for 15 weeks from both a professional language educator and native conversation teachers. Students lived in private homes and spoke only Italian while with the families.

Predictors used were 1) the five sub-scores, total score, and short-form scores from the Modern Language Aptitude Test, 2) Verbal and Math scores of the SAT, and 3) years of language training, regardless of level or type. Criteria gathered included: 1) independent proficiency ratings of reading, writing, speaking, and listening ability, and 2) a total and language grade average. In addition, Reading and Writing sub-scores from the MLA-Cooperative Foreign Language (Italian) Test were available for part of the group. Zero-order correlations ranging from .00 to .60 were found.

As expected, the MLA standardized achievement test criteria showed the highest predictability. Amount of formal language training in general, and in Italian particular showed very little relationship with any of the criteria. The MLAT short-form, requiring approximately half the administration time, as the full length form, proved to be the best single predictor. Eight variable multiple correlations ranging from .58 to .73 were also noted.

SESSION 77

Paper session --STATISTICAL METHOD AND THEORY -II

Leslie D. McLean, Ontario Institute for Studies in Education. Multiparameter Bayesian Regions.

Bayesian techniques for making inferences about the parameters of probability distributions are undergoing a period of rapid development. The best-known techniques focus on single-parameter problems and on the posterior density of the parameter given the data and some assumed (or known) prior density. Many problems in educational research, however, are necessarily multiparameter, and the joint posterior density may be the function of interest. The estimates of the parameters are often correlated and have unequal variances. Plots of joint posterior densities contain all the information in the data (and in the prior density, of course) and can reveal a variety of important relationships in the estimation situation. (See e.g. McLean, 1964; and Tiao and Zellner, 1964). Two problems suggested to be of this type are regression analysis and variance component estimation. Bayes posterior densities for these situations have been derived in general by Tiao and Zellner (1964) and Tiao and Tan (1965) as well as by others.

Applications of the general technique are shown for two educational research problems. A computer directed "plotter" is used to obtain graphs of the joint posterior distributions of up to four multiple regression weights and two variance components. Bayes posterior H. P. D. (High Probability Density) regions are calculated (Box and Tiao, 1965). Data are used from a multiple regression study of verbal and quantitative standardized tests on statistics grades and from variance component estimates calculated for reliability estimation. Box and Tiao (1965, pg. 6) make the comment with regard to the plotter output that "such plotting should be part of the normal stock in trade of the modern practicing statistician."

Ronald Thurner, University of Minnesota. Exact Values of the Power for the Test of the Homogeneity of Two Binomial Populations Using the Chi-Square Statistic with Yates Correction for Continuity.

This paper presents the exact values of the power for the test of homogeneity of two binomial populations in a 2 x 2 contingency table using the chi-square statistic with Yates correction for continuity. Tables are presented from which may be read the probability of rejecting the null hypothesis, H_0: $P_1 = P_2$, for all combinations of P_1 and P_2 from .10 to .90 in steps of .10 for the .01 and .05 nominal α levels. Sample sizes from $N_1 = N_2 = 10$ to 50 in steps of five from each of the populations are represented in the tables. The fact that the test is conservative for all sample sizes investigated is discussed.

Leonard A. Marascuilo, University of California, Berkeley. Large Sample Multiple Comparisons.

Large sample multiple comparisons based upon a χ^2 analog of Scheffe's Theorem are illustrated by means of five examples. The examples involve the correlation coefficients of K independent bivariate normal populations;

the parameters of K independent contingency tables; the parameters of K independent normal populations with unequal variances; and the differences between the parameters of K sets of paired normal populations with unequal variances. In addition, a general test statistic is presented that may be used to test the Null Hypothesis that involves the parameters.

Harry E. Anderson, Jr., University of Georgia and John L. Horn, University of Denver. The Perimetric Correlation: Explicit Functions of Geometrical Methods in the Teaching of Correlations.

The perimetric correlation is determined as a function of an areal ratio: the area, A_e, inside of the ellipse bounding the set of observation points on the scattergram in relation to the area, A_j, in the rectangle with sides defined by the ranges of the two variables. The use of such isofrequency, or equa-probability, ellipses is quite common for didactic purposes. The perimetric correlation, A_{xy}, is defined as:

$$A_{xy} = \sqrt{1-(4A_e / A_j)^2}$$

The division of A_e by A_j has the fortuitous effect of cancelling out the scaling factors, the standard deviations, and chi square. The correlations between two variables then can be computed directly as a function of the proportion of the rectangular area over which the points are scattered. A non-perimetric (principal axis) modification of Jenkins' short-cut procedure is also shown to be useful in the classroom.

SESSION 78

Paper session:--MOTIVATIONAL FACTORS IN READING

Edith M. Gifford, Edinboro State College. Test Anxiety and Reading Rate Flexibility.

Does test anxiety affect learning to vary reading speed according to purpose for reading?

For a sample of bright (average IQ 123) upper middle class fourth graders reading at or above grade level, TASC anxiety (per se and with DSC or LSC partialed out) was, except in a few instances for boys, unrelated to (1) pre-training use, (2)learning during a highly structured 10-day training session, or (3) retention or post-training improvement of rate flexibility, as measured by differentiation between reading speeds used for different purposes.

Ability to profit from training, tested by part correlation, showed significant relationship to TASC (DSC or LSC partialed out) only where boys were differentiating between reading rates for (1) remembering sequence and finding a specific fact, or (2) reading for main idea and for a specific fact. Girls' scores showed no significant relationship to anxiety. For boys, relationships between TASC anxiety alone (also when DSC and LSC were partialed out) and post-training ability to differentiate between reading speeds were all positive, hence higher anxiety appeared facilitative, though not usually significantly so. For girls, correlations were non-significant and mixed positive and negative.

Similarly, TASC anxiety was unrelated to CTMM IQ or to the STEP reading comprehension subtest, except in one instance when LSC was partialed out of the boys' TASC-STEP correlation, thus in the main disagreeing with findings published by Sarason, 1964. Boys' STEP- and IQ-TASC correlations were consistently negative. All these correlations were non-significant for girls, positive when DSC and LSC were considered and negative when correlated with TASC alone. (TASC is Test Anxiety Scale. DSC is Defensiveness Scale. LSC is Lie Scale. CTMM is California Test of Mental Maturity.)

Carolyn M. Neal, Roosevelt University. A Study of the Relationship of Personality Variables to Reading Ability in College Freshmen.

The purpose of this study was to discover the relationship between personality characteristics as measured by the MMPI and the Kuder Preference Record, and reading ability as assessed by the Cooperative English Examination and the SCAT ability test (the verbal score). In addition, the complex of the componetts of the personality variable was related to the components of the reading performance variable.

The L Scale, F Scale, and Hypochondriasis, Depression, Hysteria, Psychopathic Deviate, Pathological Sexuality, Psychasthenia, Schizophrenia, Hypomania, and Dominance scales had a weak but significantly negative relationship to reading ability.

The Paranoia, Introversion, Theoretical, Agreeable, Scientific and Literary scales had a weak, but positive relationship to reading ability.

The first factor yielded by the canonical correlation showed that the generally able reader revealed a general absence of psychopathology (except for a slight schizophrenic tendency). He tended to be introverted and agreeable, and to have literary, scientific, and theoretical interests.

The second factor showed that the fast reader with poor comprehension tended to be depressed and psychopathic. Subdivisions of the sample on the basis of sex, grade point average, and SCAT ability score revealed subsample differences in the personality patterns associated with reading ability. The results replicate those of a previous pilot study.

Anne Lasswell, Oregon State University. Goal-Setting and Performance in Self-Directed Reading Tasks.

The major purpose of the study was to investigate the influence of goal-setting upon fourth graders' performance and enjoyment of self-directed reading tasks. Ninety-six fourth graders were assigned to four training groups of 24 Ss each, stratified on the basis of sex and reading achievement level. Each group participated in a separate week-long training program designed to increase speed of comprehension in reading for the Sequence purpose under different training conditions. Group NG's members did not set goals. Group G's members set goals. Group ERG's members received information about an external reference group's scores while setting goals. Group IRG's members received information about the internal reference group's scores while setting goals.

Results revealed no significant differences among training conditions for speed of comprehension. When speed, regardless of comprehension was the performance criterion, however, significant differences were found. Subjects in Groups ERG and G spent significantly less time on first trials of selections than members of Group NG. An analysis of variance performed on goal discrepancies revealed that members of Group IRG set significantly higher goals in relation to past performance than members of Groups ERG and G. Median test analyses performed on Ss' ratings of enjoyment of the training conditions indicated that Ss in each of the three goal-setting groups enjoyed the training program to a greater extent than non-goal setters.

Samuel Weintraub, University of Chicago, and Terry Denny, Purdue University. Perceptual Consistency of First Graders toward Reading.

The authors demonstrated earlier the feasibility of categorizing first-graders' responses to questions about their perceptions of reading. Inter-rater reliability coefficient ranged from .82 to .90. This study examines the response consistency of a sample of first graders interviewed monthly throughout the school year. Eleven Ss were drawn from three first-grade classrooms in a lower middle-class neighborhood. Each month of the school year E interviewed each S to assay Ss percepts of reading. The responses to three questions were examined: "What is reading? Do you want to learn how to read? Why?"

Response categories employed in the analyses of the question "What is reading?" were (1) Do not know, no response, (2) Vague, tautological, (3) Cognitive, (4) Object-related, (5) Valuative, (6) Mechanical, (7) Expectation: Those used in the analyses of the question "Do you want to learn how to read? Why?" were (1) Do not know, no response, (2) Vague-tautological, (3) Intrinsic--self, (4)Goal seeking, (5) Affect--value statement (6) Identification, (7) Negative. Results showed that the subjects entered their first-grade classes and left them with essentially the same perception of what reading is and why they wanted to learn to read.

SESSION 79

Paper session --VERBAL LEARNING

Julius M. Sassenrath, University of California at Davis. *Meaning and Verbal Learning.*

The purpose of this experiment was to determine the effects of denotative and connotative meaning of English words on Learning a verbal paired associates task comprised of such words. From mediation theory, it was predicted that with increased denotative meaning of words and increased connotative meaning of these same words from neutral to either negative or positive on the *Semantic Differential* there should be fewer errors in learning to pair associate these words.

Phase I of the experiment involved administering an 80 word Vocabulary Inventory (denotative meaning test) selected from Irving Lorge's unpublished count of 5 million words from contemporary magazines to 172 university students. The same 80 words were rated by the same 172 *Ss* on 10 scales (4 evaluative, 3 potency, 3 activity) of the *Semantic Differential* (connotative meaning). From these data, 12 paired associates were selected; 6 high and 6 low on denotative meaning, with 2 of the 6 being positive, 2 neutral and 2 negative on connotative meaning. Phase II of the experiment was a 2 (high and low denotative meaning) by 3 (positive, neutral, and negative connotative meaning) treatments by *Ss* design in which the appropriate 2 paired associates were scored in each of the 6 cells. Thirty-seven university students, different from the initial 172, were *Ss* in Phase II.

Employing appropriate interactions as errors terms, it was found that there were significantly ($p < .01$) fewer errors in learning the high denotative than the low denotative paired associates. In addition, there were significantly ($p < .01$) fewer errors in learning the positive and negative than the neutral connotative pairs. These results can be interpreted in terms of mediation theory.

Thomas J. Shuell, and Geoffrey Keppel, University of California. *A Search for the Functional Stimulus in Serial Learning.*

Serial learning appears to be an important process in the memorization of the alphabet, the numerical system, prose passages, and other similar tasks. At first glance, it appears reasonable to assume that a serial task is learned in successive units or chunks. Paradoxically, however, attempts to isolate the functional stimulus in serial learning have not been very conclusive. While it was long assumed that each item in the serial list served as the stimulus for the immediately succeeding item, the evidence to support this assumption has not been very impressive. After discussing certain inadequacies of previous experimental designs used in testing the chaining hypothesis, the present paper presents the results of an experiment supporting the hypothesis. After learning a 12-adjective serial list to a criterion of one perfect recitation, *Ss* were given 10 trials on a double-function paired-associate list in which the successive elements of the serial list were preserved (E groups) or which consisted of an unrelated set of adjectives (C groups). Two E and C groups differed in the method of stimulus presentation during serial learning. In Cond. S the starting point remained the same on all trials, while in Cond. V the starting point was varied on each trial. Subjects were fully informed of the relationship between the two lists prior to paired-associate learning. The E groups produced marked positive transfer which was significantly greater for the V method on the first transfer trial. Plotting the mean number of correct responses for each item over the ten transfer trials as a function of the item's position in the serial list for Cond. S revealed the typical bowed curve for Grp. E but not for Grp. C. Further, there was significant positive transfer at every serial position. It was concluded that the associative strength formed between contiguous items in the serial list is the critical factor producing the positive transfer observed on the paired-associate list.

Gerald W. Faust, and Richard C. Anderson, University of Illinois. *The Effect of Incidental Material on the Learning of Russian Vocabulary from the Copy Frame.*

In its primitive form, a "copy frame" consists of the same sentence presented twice with a blank in the place of a key word in the second presentation. Such an arrangement permits the student to complete correctly the blank without attending to that part of the stimulus which it is hoped will elicit the response in the future. On the basis of the dictum that only active responses are learned, persons preparing self-instructional programs are often advised to avoid long frames containing incidental material. Incidental material may play a heretofore overlooked role; the student must find the stimulus term and discriminate it from the surrounding material in order to fill in the blank. The prediction is that, despite possible interference from associations acquired from incidental material, the addition of such incidental material to the copy frame will result in a net increment in learning by facilitating S-R hookup.

Each of two versions of a program teaching Russian vocabulary was presented to 24 summer students enrolled in an educational psychology course. Both versions of the program contained a training sequence entailing 10 presentations of 12 English-Russian word pairs, a posttest, interference material presented by the anticipation method and finally a retention test on the initially-learned Russian words. On version used only copy frames (a prompt sentence with English subject and Russian equivalent for a predicate nominative and the same sentence with the Russian word replaced by a blank) in the learning sequence. The second version was identical except that the prompt sentence was embedded in a paragraph with four other English to Russian sentences.

The group receiving the program containing the incidental material performed better than did the group receiving the copy-frame program on both the posttest and the retention test.

SESSION 79

Paul E. Johnson, University of Minnesota. *The Associative Organization among Concepts in Subject Matters: An Illustration from Physics.* The present paper is concerned with:a) the relationships among responses to items on a verbal association test when these item are words that represent concepts in a school subject matter, and b) the usefulness of viewing the totality of these relationships, or associative structure, as a map of cognitive structure in a specific subject matter area. The experiment reported in this paper involved a verbal association test in which subjects wrote all the words a given test item made them think of in a minute's time, and a rating scale test in which subjects judged pairs of the items which appeared on the association test for their degree of conceptual similarity. Measures of associative structure were correlated with performance on the rating scale test and with achievement in the subject matter.

Stanley L. Deno, Paul E. Johnson, Joseph R. Jenkins, University of Minnesota. *The Structure of Association to Words and Pictures.* The present paper is concerned with the organization among subjects' responses to items on a free association test (a) within a single set of stimuli consisting of either words or pictures, and (b) between each stimulus set. In the experiment reported in this paper, two groups of subjects gave verbal free associations to both 43 pictures and 43 words which were labels for pictures. An independent sample of subjects was used to establish the labels for the pictures. One group of subjects responded first to the pictures and then to the words, while the second group of subjects responded to the words first and then to the pictures. Associative meanings were determined for both pictures and words separately. These associative meanings were then used to determine the structure of subjects responses within each stimulus set as well as the associative overlap among each picture word pair.

SESSION 82

Paper session - SPECIFIC TEACHER BEHAVIOR

Thelma L. Baldwin, University of Wisconsin and Thomas J. Johnson, Washington University. Efficacy of Social Reinforcement as a Function of Teacher Punitiveness and Expertise.

One-hundred-sixty high school subjects in groups of five heard tape recordings of simulated classrooms in which two aspects of teacher power, punitiveness and expertise, were varied. Forty-eight of the subjects then participated in an interview with the teacher whose class they had heard. During 3/4 of the interviews, the teacher attempted to condition a verbal operant response with positive social reinforcement. 1/4 were conducted as the control, non-reinforcement condition. The major finding indicated that non-punitive teachers elicited a higher operant rate than punitive teachers when administering positive social reinforcement, and that punitive teachers elicited a higher operant rate than non-punitive teachers under the control, non-reinforcement condition. Expertise produced no significant effects on the operant rate.

Data from two questionnaires included in the study yielded strong supporting evidence for the success of the experimental manipulations of teacher punitiveness and expertise, and also for the credibility of the tape-recorded classroom simulations.

In addition to the main finding, two aspects of the study should be of interest to educational research. The use of power or social influence as a classroom variable seems a fertile area for further investigation and secondly, the successful tape-recorded classroom simulation suggests a practical methodology for careful manipulation and control of classroom variables.

Jimmie C. Fortune, Memphis State University, N. L. Gage, Stanford University and Robert E. Shutes, Palo Unified School District. The Generality of Ability to Explain.

The question of generality of ability to explain is crucial to efforts to measure, understand, and improve this function of teachers. How general is this ability, over topics, over pupils, over both topics and pupils? The ability was measured by the adjusted mean score, on a 10-item comprehension test, of a 5-pupil group to whom one of 30 social studies interns had explained, in 15 minutes, while being video-tape recorded, one of 20 Reports from the Atlantic Monthly. Each mean score of a group on a Report was adjusted for the mean on the Report of all groups of pupils and also for the mean of the group on all 20 Reports explained to it. The 30 interns were divided at random into 5 blocks of 6 each; each intern taught Report A to Group 1, Report B to Group 2, Report C to Group 1, and Report A to Group 2.

Correlations between adjusted mean comprehension (i.e., ability-to-explain) scores were computed within each of the 5 blocks ($N = 6$) and also over all blocks ($N = 30$). Results indicated substantial generality over pupils (mean $r = .4$); i.e. ability to explain Report A to Group 1 correlated significantly with ability to explain Report A to Group 2.

But there was no significant correlation between ability to explain Report A to Group 1 and ability to explain Report C to Group 1, nor between ability to explain Report A to Group 1 and ability to explain Report B to Group 2.

John A. Buelke, Western Michigan University. Responses of Learners to Teacher Explorations of their Academic Behavior Patterns

The basic context of this research is "teacher education". This study is one of several now complete as tests of the hypothesis that "the responses of individual learners to the strategies employed by teachers with them have no significance in the evaluation of either the strategy or the learning".

Several strategies have been employed experimentally by this researcher with a view to conceptualizing to prospective teachers certain dimensions of their own behaviors. In these field studies, the academi-personal behavior patterns of learners have been penetrated in ways designed to threaten their self-images. In this study, 71 learners from two university groups were: 1) confronted with a task-assignment; 2) required to appraise themselves; 3) taught in direct reference to their individual self-appraisals and the task-assignment; 4) reconfronted with the same appraisal instrument; and 5) threatened on the 4th day for inadequate performance. In general, the students behaved as: 1) objectively interested; 2) non-perceptively confused; and 3) non-teachably hostile to teacher and/or strategy.

It would appear that strategies employed by an individual teacher need to be evaluated in direct reference to the ways that academic-personal behavior patterns of learners are penetrated in threatening manners.

David Crispin, Indiana State University. Discipline Behaviors of Different Teachers.

This study tests the hypothesis that the number of acts of discipline on the part of the teacher is a function of the personality of the teacher. Two trained observers of demonstrated reliability used the Crispin System of Interaction Analysis to record the classroom behaviors of teachers and students in two public elementary schools and a public high school. All behaviors were recorded every three seconds unless there was a change of behavior or speaker, which was recorded as it occurred. If a behavior lasted longer than three seconds, it was recorded again, etc. Cnly verbal behavior was recorded. Since DISCIPLINE is one of the seven types of behavior set forth in the system, one can count the total number of such acts recorded by the observers. And since all observations covered class periods of 50 minutes, it can be determined whether there was a significant difference in the number of acts of discipline among the teachers in the various situations.

The findings indicate that in all three conditions observed, discipline behavior was a function of the teacher himself, not the situation.

SESSION 82

Daniel Tanner, City University of New York. Influencing College Student Opinion in the Classroom.

To what extent can the college instructor serve as a manipulator of student opinion? How is individual student opinion affected by classroom group membership? Eight classroom groups of college students were exposed to a kinescopic recording of a television lesson in elementary science. Immediately preceding the exposure to the kinescope, the students in each class were informed by their instructors that student ratings of the lesson would be made on a 50-item semantic differential scale. All students were informed that their ratings would be strictly anonymous. In two of the classes, the student evaluations were elicited with the instructor playing a neutral role. Prior to the elicitation of ratings in two of the other classes, the instructors offered their own opinions (in one class, the instructor's opinions were favorable toward the teacher in the film, while in the other class his opinions were negative.) An identical procedure was followed in two additional classroom groups, but in each of these instances three students in the group were enlisted beforehand to deliberately reinforce the instructor's opinions. In yet another two classes, an outside "expert" was employed to offer his opinions with the instructors playing a neutral role.

The findings revealed that the opinions of the college instructors served as a highly pervasive influence in bending student evaluations positively and negatively on the semantic differential in comparison to the ratings elicited by the control groups. When enhanced by the rigged opinions of students, the influence on group opinion was even more pervasive. The outside "expert", while exerting a significant influence on group opinion, failed to measure up to the instructor as a force in molding student opinion.

SESSION 84

Paper session --PROGRAMMED INSTRUCTION

Frederick G. Knirk, Wayne State University. An Analysis of Time and Cost Factors Involved in Using Instructional Technology Systems in the Classroom.

If the instructional technology can attain the instructional objectives desired by the teacher for his students, the time and cost requirements of the various instructional systems becomes relevant; i.e., can a televised course be afforded?; does programing a course significantly reduce the required instructional-learning time? There is much research being done which relates to the time and cost requirements of instruction by media. Usually the comparative time and cost requirements are recorded as secondary findings, if they are measured at all (especially true of cost figures).

Also included in the paper is a brief discussion of instructional systems to indicate that instructional technology cannot exist independent of a teacher. Man and machine costs must be examined and reported in instructional technology systems.

Time requirements are reported (1) by instructional media, (2) by an assessment of objective attainment, (3) by subject area, and (4) by grade level. The cost input requirements are summarized by an examination of (1) the overall expenditure of educational funds, (2) by media and the size of the using population, (3) by media averages in per student per hour units, and (4) by a detailed examination of one media form (programed instruction) by reporting the figures by subject, assessment of learning, time requirements and the cost requirements by study.

M. Daniel Smith and A. Joyce Bolyard, Earlham College. Investigating Certain Variables via a Programed Sequence of Concept Formation Experiences Related to Vector Spaces.

Learning tasks were constructed in which no words were involved. Each represented one stage in a programed sequence, and each stage was a concept formation problem. Data recorded were (1) number of exemplars to criterion, (2) mean time per response, (3) mean accuracy of response. Performance of 10 5th grade Ss was compared with that of another 10 who covered a similar sequence with word tasks added to each exemplar. Simultaneous and successive presentations were also involved. A trend analysis of (1) above indicated curves for wordless and wordal versions were significantly different (.05) and these treatments differed on (2) and (3) (.005). Simultaneous was significantly different from successive on (2) only (.01).

John M. Gordon, Jr., University of Hawaii. Interaction Effects of Varying Step Size and Feedback in Programed Instruction.

A three-way factorial experiment consisting of two feedback levels, four step size levels, three achievement levels, was conducted to estimate the interacting effects of varying program step size and feedback upon students representing different achievement levels. Over 400 seventh graders were divided into three achievement levels by their scores on a teacher-made test, and randomly assigned within levels to one of eight variations of a general science program. Each variation represented one of four arbitrarily determined step size levels, and either giving and withholding feedback. Separate knowledge and application criterion instruments were administered as well as affective rating scales. Program error rates and completion times were also noted. The cognitive tests were readministered after the summer vacation. Entry achievement level accounted for the major (90%) overall cognitive test variance. There were no feedback or step size by achievement interactions. In terms of the knowledge items, proving feedback only became facilitative with the moderately difficult step size (approximately 20-30% mean error rate). In terms of the application items, those receiving the most difficult program (40 to 80% Er.) *without* feedback did equally as well as those using the easiest (2 to 8% Er.) *with* feedback. In terms of affective ratings, those receiving feedback rated themselves as less bored. Only entry achievement level effected retention as the lowest third lost less, but they, of course, had less to lose. There was no consistent indication of any relationship existing between any of the cognitive measures; prior achievement test scores, error rate, knowledge test scores or application test scores, and any of the affective ratings.

R. J. Karraker, University of Missouri at Kansas City. *The Multiple-Choice Question and Linear Programming.*

One of Skinner's objections to branching programming is that "--- effective multiple-choice material must contain plausible wrong responses, which are out of place in the delicate process of shaping behavior because they strengthen unwanted forms." To test this hypothesis, 72 college freshmen enrolled in educational psychology were stratified into high and low ability groups on the basis of their score on the Gamma Form of the Otis Quick Scoring Mental Ability Test. Within their ability group, *Ss* were randomly assigned to one of three treatment groups.

Group 1 was administered a conventional multiple-choice test of 40 items, and was not told the correct answers. Group 2 took the same test, but was told the correct answers. Group 3 did not take the multiple-choice test. The criterion for all three groups was a recall test utilizing the same stem of the multiple-choice items, but requiring *Ss* to compose (or recall) the correct answer. Their score was the number of recall responses that were in essence the same as one of the three wrong alternatives in the multiple-choice questions.

A 3x2 treatment by levels factorial analysis of variance revealed significant differences ($<$ .01 among groups in treatments and levels. A Duncan Multiple-Range Test revealed Group 1 did significantly worse ($<$ .01) than either Groups 2 or 3. The low ability group did significantly worse ($<$ .01) than the high ability group. The interaction was not significant.

From this study it would appear that the multiple-choice question does not strengthen unwanted forms of behavior, *if Ss* are given knowledge of results.

William H. Allen, University of Southern California and Robert T. Filep, System Development Corporation. Visual and Audio Programed Instruction Presented in Teaching Machines.

This study investigated the effectiveness of three visual stimulus modes (verbal, graphic, and motion) and three audio stimulus modes (redundant), directive and silent sound) presented in machine-mediated programed instruction.

Each stimulus mode was studied in relation to programing technique (linear or branching) and to learner characteristics of I.Q., Achievement, Non-Verbal Ability, Sex, Ethnic and Parent Occupation classifications. Thirteen hundred 8th grade students enrolled in four different junior high schools of the Los Angeles City Schools comprised the experimental population.

Three parallel experiments were conducted using subject content designated as Non-Concrete, Concrete, and Action-Process. Programed instruction sequences requiring constructed responses were presented using automated teaching devices that incorporated 16 mm motion picture film and slide projectors.

The three separate experiments combined three visual presentation modes with each of three audio implementation modes and each of the programing techniques comprising a 3 x 3 x 2 factorial design. The flexibility of the teaching devices permitted testing of hypotheses about ways to present information for optimum learning, in this case some 54 hypotheses.

The relative effectiveness of the main variables and the combinations of variables in teaching, the acquisition of knowledge and the application of the learning will be determined by means of Multivariate-Covariate and Kendall-Tau analyses. In addition, response time, error type and acceptance of the mode of presentation will be determined. This study was conducted under United States Office of Education Cooperative Research Project No. 1956.

SESSION 85

Paper session --PERCEPTUAL AND MOTOR SKILL LEARNING

David B. Crr, and Herbert L. Friedman, American Institute for Research. Recent Developments in the Comprehension of Time Compressed Speech.

Past research on faster than normal speech published in the Journal of Educational Psychology has demonstrated two major findings: (1) that naive listeners (untrained) can comprehend modestly speeded material at approximately twice normal presentation rates with comparatively little loss of comprehension (10% to 20% losses compared with normal speed comprehension); (2) that with 10 to 15 hours of spaced practice, comprehension levels at higher rates of speed (for example 425 words per minute) show statistically significant improvement.

Further recently completed work is presented. Seven subjects were exposed to practice material consisting of time-compressed talking books at 425 words per minute, for a period of ten to twelve hours a day for five days. A bench-mark comprehension test was given at the end of each days practice. A sixth test was given at 475 words per minute before the experiment began and repeated again at 475 words per minute at the end of the experiment. The findings showed that the subjects averaged approximately 20% of normal speed comprehension for the pre-experimental administration of the 475 words per minute passages, which value rose to 52% at the post-experimental administration. Scores on the benchmark passages averaged 40% at 425 words per minute at the end of the first day and rose to 70% at the end of the fifth day.

Leonard J. West, City University of New York. Vision and Kinesthesis in the Acquisition of Typing Skill.

The prevailing but hitherto untested hypothesis about the acquisition of perceptual-motor skills is that visual cues are important at the start, but that as performance becomes habitual, proprioceptive feedback or "feel" becomes the more important. To test this hypothesis, 263 typists ranging in skill from 9 to 108 word per minute speeds worked under each of three conditions: (1) under normal conditions, furnishing a base measure of skill, (2) under instructions to retype in the case of error, before continuing, and (3) under the same instructions but deprived, by means of a paperboard shield, of visual reference to the typewriter or to their typescript. Awareness of errors, as indicated by retyping is the evidence for dependable feedback; Condition 2 furnishing a measure of all-senses feedback; Condition 3, of kinesthetic (nonvisual) feedback. Three sets of copy materials of equal difficulty were used for the three work conditions, in counterbalanced order.

Analysis of data showed that insistence on "touch" typewriting from the start of learning, beginners have very little dependable kinesthetic feedback (20%) and that even experts (95+ wpm) sense their errors, when working without vision, only 50 percent of the time. These and ancillary data (e.g. on the effects of visual deprivation on performance) tend to support the hypothesis.

SESSION 85

Orpha K. Duell, and Richard C. Anderson, University of Illinois. Pitch Discrimination Among Primary School Children.

One-hundred-sixty-two first, second, and third graders were tested to determine the smallest interval Ss of this age can discriminate. The Ss were presented pairs of pure tones which had been recorded on tape. They judged the second tone as higher, the same, or lower. The standard ranged from 390 to 440 cps. The intervals ranged from 1/3 of a half step to a major sixth both above and below the standard. The curves for the second and third grades sloped steeply from the 1/3 of a half step to a major second and then continued to rise gradually; however, the curve for the first graders showed little variation from chance for all intervals.

In a second experiment, 168 Ss of the same age group were tested. Some changes were made in the method to correct difficulties revealed in the first study. The Ss now judged the second tone as the same or different. The results (see Fig. 1) show that the majority of these Ss can discriminate intervals as large as and larger than a half step; however, about 4% are unable to discriminate differences as large as a sixth. Though the differences in performance between grades are significant ($p < .01$), the shape of the performance curve is now the same for all grades. Within a 30 minute testing period, children's pitch discrimination performance does not change.

Robert B. Hayes, Department of Public Instruction, Harrisburg. Immediate Learning Reinforcement.

This two-year study was an attempt to determine the effectiveness of immediate standardized standardized reinforcement in learning a complex mental-motor skill. An electronic educational laboratory with 12 simulated cars flashed danger red lights to tell each learner when an error had been committed in one of the following areas: speed, steering, brakes, turn signals, or headlights. The learners in the simulator reacted to a series of movies programmed from the easy to the more difficult tasks of learning to drive a car.

Each year of this study, 400 high school students were randomly assigned to either a movie-simulator centered treatment or a dual-control car program. The pre-experimental variables were sex, I.Q., age, grade average for prior school year, driving knowledge, and attitude toward driving. The criterion measures were: (1) driving skills measured by two independent judges during the fourth and final periods of practice driving; (2) the number of times to pass the State license examination; (3) the number of accidents and traffic violations during the first year of driving; and (4) post-test results for attitude and knowledge. Statistical methods included chi-square, analysis of variance and covariance.

The results indicated that with an integrated teaching approach, a simulator with immediate learning reinforcement can be used to substitute nine hours of simulation for three of six hours of behind-the-wheel dual control car training. It appears that with efficient scheduling to take advantage of the improved student-teacher ratio, the simulator could be used to greatly increase the number of students trained per instructor each year.

SESSION 87

Paper session --MEANINGFUL LEARNING AND TRANSFER

R. L. R. Overing, University of Utah and R. M. W. Travers, Western Michigan University. *An Investigation of the Effects of Irrelevant Information upon Transfer with Preadolescents.*

The authors refining and extending the work of Judd (1908) and Hendrickson and Schroeder(1941) investigated the effect upon transfer of varying irrelevant information whilst holding relevant information and time constant. Reviews of the concept formation and other related literature suggested that visually compressed material, e.g., line drawings, would produce more learning than realistic conditions which possess much irrelevant information, but Travers *et. al.* (1964) questioned whether these findings would also hold true for transfer situations. In Exp. 1, two groups of 24 *Ss* ($M = 12$; $F = 12$; Grade 6; I.Q. Range 100-130) each received;a) instruction in the principle of refraction, and b) a demonstration of the visual displacement of an underwater object. Group 1 received instruction using three-dimensional physical apparatus, Group 2, using two-dimensional large scale, colored line drawings. The verbal component of instruction for both groups was held constant. Transfer tasks involved aiming correctly so as to hit an underwater target. *Ss* who had learned in the presence of irrelevant information performed significantly better ($P < 01$, Duncan's New Multiple Range Test) than *Ss* taught under conditions designed to minimize irrelevant information. The authors interpret the results as suggesting that training with compressed information fails to give the learner experience in dealing with situations involving many irrelevant cues.

E. Z. Rothkopf and E. E. Bisbicos, Bell Telephone Laboratories. *Learning from Written Instructional Materials: Selective Facilitative Effects of Test-Like Events.*

Several experimental investigations have shown that test-like events such as periodic questioning of studying students increased how much they learned from written instructional materials. Recent evidence indicates that facilitative effects of questions have at least two components: a) the direct instructive effect, i.e., the questions contain informative material, and b) a general set-like effect which resembles attention, i.e. , questions affect the studying behavior of students.

The objective of the present experiment was to find out whether : a) the set-like effects resulting from questions were general or whether they could be attuned to particular attributes or subsets of the written instructional text;and b) do these set-like effects change during the course of study.

High school students ($N = 324$) studied a 9000 word factual passage which was divided in 12 three-page segments. Nine experimental treatments differed with respect to a) location of adjunct questions relative to the appropriate segment, and b) homogeneity of the conceptual class of the correct answers to questions. The conceptual classes included quantities, proper names, technical phrases, and common phrases.

Treatments were compared on a test administered after the end of the study.

Results indicate that the character of the questions can have a marked facilitative effect (tuning) on learning from subsets of the text. Progressive increases in "tuning" as a function of exposure to questions and text were observed. The findings confirm that it is useful to conceptualize the effect of adjunct questions on learning from written material in terms of changes in mathemagenic behavior.

Joseph M. Scandura, Florida State University. The Basic Unit in Meaningful Learning - Association or Principle.

A scientific language (SFL) will be introduced for formulating research problems on complex learning and teaching. In the SFL, the principle, rather than the S-R association, is the basic behavioral unit. More specifically, the SFL and S-R mediation descriptions of classical learning types will be contrasted, the former using mathematical symbolism (i.e. of a mathematical function) and the latter using traditional S-R schema.

The following advantages of the SFL will be discussed: 1) The SFL provides a symbolic means for representing all classical learning types, from simple association to higher order principle. No S-R mediation representation of principle learning has been advanced to the author's knowledge; 2) The SFL does not become cumbersome as situations increase in complexity. The predictive value of the S-R language typically becomes less precise as situations become more complex. Many subject matter complexities can be dealt with precisely in the SFL. In particular, the set-function language can be used to formalize many research problems, of interest to educational psychologists, which have not lent themselves to S-R analyses. Comparisons of expository and discovery modes of instruction and the conservation of quantity are but two of these areas; 3) The SFL provides a basis for pointedly discussing the important question, "what is learned". By identifying the important stimulus dimensions prior to learning, it is possible - under certain conditions - to determine whether a principle has been learned or not by presenting a test stimulus and observing the response. Data on this point is available. Further implications of the SFL and relevant data will be discussed as time permits.

T. Antoinette Ryan, Oregon State University. Effect of Simulated-Situation Problem-Solving Tasks in Increasing Ss' Ability to Apply Principles in Realistic Settings.

The study aimed to test the effectiveness of using planned simulated-situation problem-solving tasks under conditions of varied instructional approach (T. A. Ryan, 1965) to increase Ss' proficiency in applying principles in realistic settings.

The sample consisted of Ss from the population of students enrolled in undergraduate educational psychology. Ss were assigned randomly to three instructional conditions: 1) Single instructional approach with project assignments (Control 1); 2) Varied instructional approach with project assignments (Control 2); and 3) Varied instructional approach with simulated-situation problem solving tasks (Experimental). Content, pre-treatment knowledge, instructor, time were controlled. Treatment lasted 8 weeks. Criterion measures were given to test acquisition of knowledge and proficiency in applying principles in realistic situations.

Analysis of variance across treatments revealed significant differences between control and experimental conditions. The study supports the hypothesis that Ss' ability to apply principles in realistic situations increases under instructional conditions using planned sequence of simulated-situation problem-solving tasks and varied instructional approach allowing student choice of teaching method. The study clearly demonstrates that effectiveness of instruction is related directly to amount of experience-creating and environment-manipulating that the teacher does - rather than extent of information-giving.

M. David Merrill, George Peabody College for Teachers and Lawrence M. Stolurow, University of Illinois. *Hierarchical Preview vs. Problem Oriented Review in Learning an Imaginary Science.*

Seventy-one university freshmen men were presented the imaginary Science of Xenograde Systems (Merrill, 1965) by means of a teaching program in which the main line frames presented no informational content but merely posed a question. All of the information was presented in the summary-preview and/or the review frames. The material was presented on the SOCRATES computer based teaching system. The *S*s were divided into six aptitude combinations on verbal and quantitative SCAT scores, and *S*s in each condition were randomly assigned to 6 treatments conditions. *S*s in Group 1 received only the questions, *S*s in Group 2 received a hierarchical summary prior to the questions, *S*s in Group 3 were presented relevant summary statements (general review) after each question they missed, *S*s in Group 4 were shown how to work the problem step by step (specific review) whenever an error was made, *S*s in Group 5 were given the general review after their first error and the specific review after a 2nd error, *S*s in Group 6 were presented the hierarchical summary prior to the questions and the two stage review like group 5 whenever they made errors.

The results were: 1) Presenting the summary before the problems does *not* take significantly more time than when the summary is not presented, but it *does* increase the number of correct responses in the learning session and on the test; 2) Presenting the summary statements specific to and following each incorrect answer takes more time than not receiving them, but does *not* increase significantly the number of correct responses during learning or on the test; 3) Presenting a step by step solution following each incorrect answer *does* take more time than not receiving this review, but it also increases the number of correct responses during learning and on the test; 4) Differences in verbal and quantitative aptitude did not make a significant difference in the speed with which the task was learned or in the performance during learning and testing.

SESSION 88

Paper session --MEASUREMENT IN LANGUAGE ARTS AND READING

Douglas Porter, Harvard University and the Job Corps. Design and Preliminary Validation of a Brief, Wide-range Reading Test for Use in the Job Corps

There has long been need for reading tests with the following characteristics: 1) Wide range, covering reading attainment all the way from the early primary grades to the level of the average literate citizen; 2) Having content that is not offensive to an adult or teen-ager but still conceptually simple enough to use with children; 3) Brief enough to be administered in the wide variety of circumstances where typical reading achievement tests are so long or complex that they discourage the subjects or are impractical; 4) Bearing some specified relationship to a corpus of reading matter so that the resulting score can provide a useful criterion-oriented interpretation rather than a grade-equivalent score of little practical use in the assignment of instruction.

The Job Corps Reading Test, RJSI, was designed to meet the above specifications, to be inexpensive, and either machine or hand scorable. The following aspects of the reading test are described: 1) The corpus of reading matter from which items were generated. 2) Procedures for producing items. 3) The multiple-contingency criteria applied to each item in the process of item analyses. 4) Performance data to date, including: preliminary norms, discrimination indices, and reliability.

Suggestions are given for the application of similar test construction techniques to other instruments.

William R. Powell, University of Illinois. A Validity Check of Selected Reading Diagnostic Tests

This study offers comparative validity evidence for five reading diagnostic tests using as criteria a standardized reading test, teacher ratings, and material from the records of a small sample of fourth grade pupils. The following reading diagnostic tests were administered: 1) Bond-Clymer-Hoyt Silent Diagnostic Reading Test; 2) Doren Diagnostic Reading Test; 3) McKee Inventory of Phonetic Skill; 4) McCullough Word Analysis Test; and 5) Roswell-Chall Diagnostic Reading Test.

Simple and multiple correlations were applied to the data collected in this study. Intercorrelations of the criteria were performed; simple correlations were applied between the diagnostic tests and the criteria; total scores were correlated with the criteria; and lastly, multiple correlations were performed.

Correlations between the predictors and the standardized reading test criterion were the lowest. Further, it was noted that some of the tests predicted the criterion better than others, and certain sub-tests seemed to relate to general reading ability (or the lack of ability) better than other variables, indicating that they may be more important components of reading skills.

In general, the validity data found in this study appears to be in opposition to the sketchy validity figures offered by the test's authors. Although in most cases, multiple correlations indicated that the diagnostic reading tests did have a great deal of latent predictive power.

G. Phillip Cartwright, University of Pittsburgh, *Techniques of Analysis of Written Language.*

This paper will examine a number of techniques which have been used by researchers in an attempt to obtain objective and quantifiable evaluations of written compositions. The techniques are grouped into the following categories:

1. *Grammar.* Here are included such analyses of compositions as type and frequency of grammatical errors, knowledge of grammar, and ability to judge correctness of sentences.
2. *Sentence Analyses.* Sentence length, frequency of usage of various types of sentences, subordination, and other sentence complexity indices will be discussed.
3. *Parts of Speech.* Frequency of usage of parts of speech and numerous ratios involving different parts of speech have been developed by both traditional grammarians and structural linguists.
4. *Linguistic Measures.* These include such techniques as the type-token ratio, Yule's characteristic "K", and Zipf's wave-length.
5. *Rating Methods.* The contribution of structured rating techniques, including those stemming from factor analytic studies will be discussed.

Representative studies in which each technique has been used will be reviewed in terms of the effectiveness of the method. The problem of criteria by which to judge the written compositions will be the central theme of the discussion of the techeiques.

Paul Weener, Loren Barritt, and Melvyn T. Semmel, The University of Michigan. *A Critical Evaluation of the Illinois Test of Psycholinguistic Abilities.*

The Illinois Test of Psycholinguistic Abilities (IPTA) was developed by Kirk and McCarthy, to identify "psycholinguistic abilities and disabilities in children between the ages of 2-6 and 9." (Kirk and McCarthy, 1961). The ITPA has been widely used for research and diagnosis. Little information is available about the test except for articles and monographs by its authors. It is the purpose of this report to analyze the procedures used in the development of the ITPA. An outline of the major sections of the report and general conclusions follow:

The Model: An adaption from Osgood's Extension of Hull's mediation hypothesis. A tabular display of the relationship between model and tests is presented, along with a description of each subtest.

Standardization: Adequacy of norming sample is discussed. Implications of a restricted norm group for test interpretation are made explicit.

Reliability: Internal consistency and test-retest stability estimates of reliability are discussed. Internal consistency coefficients indicate homogeneity of variance on the subscales. Stability coefficients for subscales are low.

Validity: Concurrent, predictive, content, and construct validity are discussed. Concurrent and predictive validity coefficients are low. Internal consistency coefficients for subtests are used to indicate content validity.

Factor analysis of subscale scores indicates an overlapping structure. Empirical support for the construct validity of the test is weak.

Blanche L. Serwer and Albert J. Harris, Division of Teacher Education, The City University of New York. How First-Grade Teachers Spend Their Time When Teaching Language Arts.

The 48 first-grade teachers in the CRAFT Project, a project supported by U.S.O.E., maintained time logs for five consecutive days for each of four months. Each of four teaching methods was employed by 12 teachers. Data will be presented on average language arts instructional time and time for each of 11 kinds of language arts activities, for each of the four teaching variables. Correlations with achievement tests given in May, 1965 will also be presented and analyzed.

SESSION 90

Paper session -- TEACHER TRAINING AND EVALUATION

William H. Creswell, Jr., J. Thomas Hastings, and Warren J. Huffman, University of Illinois. *School Health Education Study: An Evaluative Report.*

The central purpose of the School Health Education Study, which began in September 1961 is to effect a long range program of improved health instruction in the elementary and secondary schools of the United States. Using a concept approach, experimental teaching-learning guides were developed and tested in four tryout school centers.

The objectives of the new health education curriculum materials were to: 1) effect a greater student knowledge than that attained by means of the traditional curriculum; 2) reveal inter-relationships of health knowledge facilitating student understanding; 3) produce increased student interest in the study of health; 4) improve teacher effectiveness and interest in health instruction. A variety of evaluative techniques were employed including questionnaires, interviews, teacher logs, and class observations.

This paper will focus on the research procedures involving experimental and control classes established in the tryout centers. The research design was adopted for the purpose of assessing the effects of the new curriculum materials. Pre and post test score measures were obtained on students in the experimental and control classes by use of attitude scales, standardized health knowledge tests, and specially designed achievement tests. Gain score differences between groups were statistically evaluated. A further step in item analysis of the achievement tests was conducted to determine effectiveness of specific elements of the new curriculum materials.

W. Owen Scott and Henry Miller, The University of Georgia. *The Professional Training, Specific Teaching Assignments, and Job Satisfactions of Georgia Social Studies Teachers.*

Using a random sample of one-third of the secondary schools (schools containing Grade Seven and above) in the state, the authors ascertained the certification status and specific teaching assignments of teachers of the social studies. In addition, information was obtained concerning their job satisfactions. More than 90% of these teachers were professionally certified to teach the social studies, while about 1% held emergency certificates. About 90% teach one or two different social studies subjects, while about 1/2 of 1% teach as many as four. The pattern of teaching assignments in the state was found quite similar to those in other southern states and over the United States as a whole. With respect to job satisfaction, helping students learn and "seeing" them progress through school was reported most frequently as the source of most satisfaction, while non-teaching duties, controlling pupil behavior, and coping with pupil heterogeneity were reported most frequently as being the least satisfying aspects of teaching. The specific patterns of certification and of teaching assignments are elaborated upon in the full report as is the information concerning job satisfaction.

Isobel L. Pfeiffer, Kent State University. Comparison of Verbal Interaction and Goals of Teachers Teaching Classes of Different Ability in Eleventh Grade English.

This study investigated the following questions in situations where each teacher of 11th grade English was teaching one high ability class and one average class or one low ability class. a) Does the teacher's pattern of classroom verbal interaction differ between the classes? b) Do the teacher's cognitive goals indicated in an interview differ between classes? c) Do the cognitive goals inferred from test items differ between classes? d) Are the teacher's cognitive goals expressed in an interview different from those implied by test items for the same class? Eleventh grade English teachers who were teaching English classes on 2 different ability levels in a large suburban high school were observed for 6 sequential class meetings. Flanders' Interaction Analysis system was used by trained observers. Individual and summed matrices for the 2 classes of each teacher were compared. Teacher-made examinations for each class in the study were analyzed according to the major categories of Bloom's Taxonomy: knowledge, comprehension, application, analysis, synthesis and evaluation. Additional data about the teachers were obtained through depth interviews with the teachers. Comparisons were made between the cognitive goals indicated in the test item analysis and those stated by the teacher. Also compared were the goals for the classes of different ability taught by the same teachers. Results raise serious questions concerning widely held views about what teachers do and how they differentiate instruction when teaching classes of differing ability. These findings are discussed with respect to organization for teaching and instructional improvement.

James B. Macdonald, University of Wisconsin. A Research Oriented Elementary Education Student Teaching Program.

Major questions asked were: 1) Do students in a research oriented student teaching experience show significantly greater gains in general teacher education outcomes? 2) What is the relationship of general outcomes of a teacher education program to student social and personal characteristics? Three groups of 22 student teachers were assigned randomly to experimental (research oriented), experimental-control (modified programs, but no research) and comparison (regular program) conditions. The experimental group received research training and did classroom research studies during student teaching. Data were collected on 16 general outcomes in cognitive, affective and performance areas. Further data reflecting personality characteristics, attitudes and values were gathered. Discriminant Analysis and analysis of covariance were used to answer the first major questions, and missing data correlation utilized to answer the second question. Results showed that the group in the experimental condition scored significantly higher than the other two groups. Correlational data revealed cognitive outcomes were more highly related to intelligence and prior achievement, and affective and performance outcomes more highly related to attitudes, values and personality variables.

SESSION 91

Paper session:--INSTRUCTIONAL OBJECTIVES

Howard J. Sullivan, Richard E. Schutz, Robert L. Baker, and Vernon S. Gerlach, Arizona State University. Instructional Components for Behavioral Objectives.

The attainment of instructional objectives specified in terms of observable learner behavior is of major concern to both teachers and educational researchers. The task of achieving such objectives would often be greatly facilitated by the use of a generic model that clearly specifies the instructional components required for acquisition of the behavior of interest in the given learning task. The purpose of this paper is to present a functional model that specifies the instructional components essential to achievement of the behavioral objectives for any learning task.

The instructional specification (IS) was developed by researchers in the Classroom Learning Laboratory, Arizona State University to provide an internally consistent model specifying the instructional and eliciting stimuli to be presented to the learner to insure his acquisition of the objective of interest. This paper reports the rationale underlying the IS and provides cues for the development and utilization of the IS in the instructional program. Also presented are relevant data from research projects employing the IS and model instructional specifications from instructional units prepared for use in elementary school arithmetic and science programs.

John M. Gordon, Jr., University of Hawaii. Some Aids in Helping Faculty Members Specify Objectives.

The advent and emphasis upon the "terminal behavior" in instruction has shed new light on the role of the evaluation specialist. No longer is the developing and scoring of tests a last-minute activity usually carried in between halves of the Bowl Games and other advantageous times. It is now (supposed to be) a companion activity to the "specification of terminal behaviors" or as was known in the good old days "determining objectives".

A recent pilot study to determine the prizes and pitfalls within the typical encounter between faculty member and evaluation specialist was undertaken at Michigan State University as one part of a larger study of media implementation in course development. Audio tapes were used to record each of four sets of meetings between a series of two specialists and four faculty members representing two different courses. Generalizations were drawn that have relevance for all who engage in this complex problem-solving task. These generalizations include suggestions as to a common terminology, where to begin, how to cope with defensive behaviors, and how to aid in making instructional decisions as well as those of item development, etc.

SESSION 91

John D. McNeil, University of California. Concomitants of Using Behavioral Objectives in the Assessment of Teacher Effectiveness.

What are the effects of using operational definitions of instructional goals in the supervisory process with teachers? Is there greater gain in pupil achievement? What are the consequences of procedures whereby teachers and supervisors agree in advance upon the indicators of successful teaching? What special difficulties arise when supervisors and teachers attempt to define operationally their intents? What are some prerequisites to the task of defining instructional outcomes to be sought?

A number of empirical studies utilizing student teachers and supervisors have been underway at UCLA the past two years in an effort to throw light upon answers to questions such as those mentioned above. These procedures include experimental studies in which teachers are randomly assigned to treatments with and without "supervision by objectives". Measures of changes in teacher behavior and pupil behavior under the two conditions have been made.

In addition, the paper will include descriptions of results from training sessions involving approximately 90 teachers and supervisors. These teachers and supervisors dealt with the task of stating objectives behaviorally and the task of using these objectives in describing and evaluating teaching. Prerequisites to these tasks have been suggested from the data obtained through these case studies. Methods for dealing with special problems in shaping supervisors' and teachers' behaviors toward objectives will be briefly described.

Clifford B. Elliott, C. L. Davis, Jr., and Philip R. Merrifield, Kent State University. Cognitive Dimensions of Lesson Objectives Set by Secondary Student Teachers.

This study was undertaken to investigate the feasibility of developing a practical procedure which would enable practicing teachers to utilize the Taxonomy of Educational Objectives: Cognitive Domain in the evaluation of lesson plan objectives. The objectives of the study were: (a) to analyze the stated-lesson-plan objectives of student teachers as recorded in their sequential lesson reports by use of the Taxonomy of Educational Objectives: Cognitive Domain: (b) to investigate whether judges can independently categorize student teachers' lesson objectives using a prescribed graphic scaling procedure based on the Taxonomy; and (c) to determine which of two specially constructed graphic scaling formats was more practicable.

Sets of four sequential lesson reports for each of twenty-five secondary student teachers were studied. Two sets of judges, each set using a different rating format, independently classified each objective according to the six major taxonomic categories.

Results of data analysis did not support the hypothesized inter-rater agreement and reliability. There was indication that slightly better results were obtained with one graphic rating format than the other.

SESSION 92

Paper session-- FACTORS IN SCHOOL ACHIEVEMENT

Richard E. Gordon, Wagner College and Richard H. Lindeman, Columbia University. Some Psychological and Biochemical Correlates of College Achievement.

This paper reports the results of studies of the relationship of some non-cognitive characteristics to the academic achievement, health, and social behavior of students in three freshman classes at a private liberal arts college in the New York City area.

Biographical and psychological data were obtained for more than 1200 students at the beginning of their freshman year, and grade-point averages and other performance data were obtained during the course of the year. Measures of blood uric acid level, triglyceride level, and cholesterol level were also obtained for a smaller sample of students from two of the three classes. Preliminary analyses showed that: (1) a number of personal and study habits reported by students, such as completing assignments on time, selecting friends having superior ability, not smoking, and rising before 7:30 AM, are related to academic achievement; (2) home conditions and other background data reported by the student, such as grandparents' birthplace, religious preference, health of student and family, and problems of personal appearance (skin diseases, wearing glasses, etc.), are related to academic achievement; and (3) certain blood chemistry measurements, i.e. uric acid level, triglyceride level, and cholesterol stress response, are related to academic achievement.

Maria Rizzo, and John M. Newell, Tufts University. Level of Aspiration, Aptitude and Academic Performance.

The present study is concerned with two groups of college freshmen who, on the basis of SAT scores and high school averages, were classified as "high risk" and "low risk" admission candidates. Degree of risk is reflected in a "predicted grade-point average" which is an estimate of the level of performance the student will attain at college. The question posed was whether the "high risk" students would be less realistic in level of aspiration than "low risk" students.

Subjects were 55 males and 35 females in each group. Two level of aspiration measures were obtained, one prior to beginning classes in September and the other in February. Predicted and actual grade-point averages, SAT scores and proposed college major were known.

The results indicate that the "high risk" students were significantly more unrealistic in the levels of aspiration than were the "low risk" students. The "low risk" students tended to raise their level of aspiration while the "high risk" students tended to lower their aspirations. In both groups, the relationship between academic achievement and level of aspiration was more realistic (i.e. significantly higher correlations). Analysis of the data by "science" versus "non-science" majors yields no differences not found with the total group.

SESSION 92

Howard R. Kight, State University of New York at Buffalo and Julius M. Sassenrath, University of California at Davis. *The Influence of Achievement Motivation and Test Anxiety on Performance in Programmed Instruction.*

This study was concerned with the influence of achievement motivation and test anxiety on performance in programmed instruction.

Undergraduates (139) were administered the Iowa Picture Interpretation (achievement imagery) test, the Sarason Test Anxiety Questionnaire, and a pretest of concepts relating to the programmed instruction. The programmed material consisted of two linear-type booklets covering fundamental aspects of test construction and evaluation. Pupil performance was analyzed using as criterion measures: (a) time needed to complete the programmed material, (b) number of incorrect responses, and (c) a retention test, administered five days after completing the material.

The analysis of variance of the data showed that high achievement motivated *Ss* performed significantly better than low achievement motivated *Ss* on all three performance criteria. Significant differences were also found between high and low test anxiety groups on two of the three performance measures, time and number of incorrect responses. High test anxiety anxiety thus appeared to have a facilitating effect upon certain aspects of performance on programmed instruction. The fact that high test anxiety groups did not perform significantly better than low TA groups on the retention test is in agreement with the hypothesis and with previous research.

Contrary to expectations, two significant interactions occurred involving test anxiety. The interaction between TA and pretest ability, with regard to the number of incorrect responses, raises some doubt as to the generalized effect of test anxiety.

One important conclusion which has stemmed from the analysis of both time and error scores is that on well structured but relatively easy material such as programmed instruction, the highly motivated *Ss* tend to work much more efficiently than the less motivated *Ss*.

Philip W. Jackson, and Henriette M. Lahaderne, University of Chicago. *Scholastic Success and Attitude Toward School in a Population of Sixth Graders*

The entire sixth grade population (11 classes located in 6 schools; N=148 boys, 144 girls) of a suburban community responded in the Spring of the year to modified versions of the Student Opinion Poll (Jackson and Getzels, 1959) and the Michigan Student Attitude Inventory (Flanders, 1965), two attitude questionnaires designed to assess the students' satisfaction with their school and their teacher. In each class, the teacher also estimated the general level of each of his students' overall satisfaction with school. Tardiness records, course grades and achievement test scores were obtained for each student.

The most striking findings were the very low correlations between the satisfaction scores and the school performance information. The correlations between satisfaction scores and teachers' grades ranged from .00 to .19; those between satisfaction scores and achievement test performance ranged from -.07 to .14. There were no significant sex differences. The teachers' estimates of the students' satisfaction were related more closely to the students' academic performance ($r \simeq .4$) than to their expressed attitudes toward school ($r \simeq .25$).

The two attitude questionnaires were moderately correlated with each other (.64 for boys; .53 for girls), and performance on the Student Opinion Poll yielded a low correlation ($r = .21$) with tardiness records, but only for boys. Differences were found among the eleven classrooms in mean scores on both of the satisfaction measures, although the rankings of the classes on the two instruments were only slightly similar (Rho = .33).

Theoretical and practical questions arising from the weak relationship between expressions of satisfaction with school and measures of scholastic success are discussed.

William L. Franzen, University of Toledo. A Cross-cultural Investigation of Arithmetic, Reading, and English-usage Achievement of Selected United Kingdom and U.S. Public School Pupils.

1,611 randomly selected United Kingdom and U.S. Public School first and second-grade and third through sixth-grade Ss of high, average and low ability, as determined by the Otis (Short Form) Ability Test, were administered the California Achievement Test Battery, modified to eliminate culturally biased items. A factorial analysis of variance was used to asses significant differences in mean achievement in the areas of arithmetic, reading and English-usage. Significant differences obtained on the main effects of grade and country and ability, and in some instances, on varied interactions of these main effects. On every variable-reading vocabulary and comprehension; arithmetic reasoning and fundamentals; and mechanics of English and spelling - significant differences were obtained on the main effects of grade and ability. United Kingdom main effect means were significantly higher than U.S. for grades one and two on five of six variables, exception being arithmetic reasoning, wherein no significant difference obtained. At grades three to six the United Kingdom means were significantly higher on variables of reading vocabulary, arithmetic fundamentals and spelling. Only on the variable of reading comprehension did the U.S. obtain significantly higher mean for grades three-to-six. On the variables of arithmetic reasoning and mechanics of English, no significant differences occurred. Inspection of means indicated that British Ss obtained higher levels of achievement in first four grade levels in nearly all subject-matter areas; by grades five and six, U.S. Ss were achieving at approximately same levels in nearly all areas; British Ss achievement revealed greater margin between high and average ability groupings than U.S. sample.

SESSION 93

Paper session --RESEARCH ON TEACHING

Fred N. Kerlinger, New York University. *The First- and Second-Order Factor Structures of Attitudes Toward Education.*

To determine the factorial nature of attitudes toward education, a 46-item summated-rating scale was constructed on the basis of previous studies and administered to three samples of graduate students of education and teachers in New York ($N = 344$), North Carolina ($N = 404$), and Texas ($N = 556$). The scale contained 23 of each of two types of items, *A* ("Progressivism") and *B* ("Traditionalism"), the reliabilities of which were found to be between .79 and .86 in the three samples. The 46 items were intercorrelated and factor analyzed with the principal factors method. Four factors were rotated orthogonally (Varimax) to test factorial invariance in the three samples; eight factors were rotated obliquely (Proequamax) for factor interpretation and for second-order analysis. The intercorrelations of the oblique primary factors were factor analyzed, and the resulting second-order factors were rotated orthogonally to simple structure.

The statistics and factor structures of the three samples were substantially similar, and the New York sample data were used in subsequent interpretation. In the eight-factor solution, five factors had *A* items with significant loadings and the other three factors *B* items. The *A* factors were tentatively labeled "Experimentalism", "Reconstructionism", "Life Adjustment", Children's Needs", and "Academic Freedom". The *B* factors were "Criticism of the Schools", "Learning as Knowledge Storehouse", and "Educational Conservatism". The second-order analysis yielded two factors, with the *A* factors loaded on one second-order factor and the *B* factors loaded on the other. They were named "Progressivism" and "Traditionalism".

Glen E. Robinson and Richard E. Scott, Research Division, National Education Association. *A Comparative Analysis of Some Characteristics of Respondent and Hypothetical Nonrespondent Groups in Nationwide Probability Samples of Teachers.*

The purpose of this investigation was to study the characteristics of certain groups of teachers who might be considered typical of nonrespondents to certain types of mailed surveys of teacher opinion.

Data obtained from NEA research Division Teacher Opinion Polls conducted in 1963, 1964, and 1965 were used in this study. The average percent of response to these polls, after adjustment for sampling frame deficiencies, was 95.7 percent. Assuming that the return approximates a complete return to each of the three samples of teachers, the investigators analyzed the comparative characteristics of five sets of hypothetical respondent-nonrespondent groupings. Criterion for grouping was the length of time the respondents had the questions in their possession; the following early-late answer groupings were used: 50-50%, 60-40%, 70-30%, 80-20% 90-10%. The characteristics of the early-late groups might be considered as indicative of the characteristics of corresponding response-nonresponse groups in similar surveys that achieve lower rates of response. Respondent characteristics selected for study were the size of the school and school system in which the respondent taught sex, marital status, age, highest degree held, teaching experience, salary, grade level taught, and geographic region.

A preliminary examination of the data indicates that there are statistically significant differences between the distribution of some characteristics for certain hypothetical respondent-nonrespondent groupings.

W. James Popham, and Eva L. Baker, University of California, Los Angeles. A Performance Test of Teaching Effectiveness.

One of the prime difficulties in developing suitable indices of teacher competence is that instructional objectives often vary in different teaching situations. Consequently, it is unreasonable to compare the performance of two different teachers when they are attempting to accomplish markedly disparate goals.

In an effort to circumvent this dilemna, the writers are developing a series of performance tests of instructor competence. These tests consist of 1) a set of operational instructional objectives, 2) a collection of activities which the teacher may wish to employ, and 3) a pre- and post-test not seen or administered by the teacher, which adhere closely to the operational objectives. The objectives and suggested activities are given to the teacher one week in advance of instruction, and he is told to prepare plans for three weeks of teaching. His pupils are both pre- and post-tested by the project staff. By stipulating identical objectives to be achieved and tested but permitting teacher divergence in accomplishing these ends, a method of comparing teacher performance without impinging on pedagogical style is provided. Teachers are compared on the basis of their pupils' achievement rather than on other, more idiosyncratic criteria. Two of the performance tests in vocational education will be available for trial in January 1966 with approximately ten teachers and 300 pupils in each subject field.

Because the major problem with such a performance test will inevitably reside in the measurement of pupil achievement and associated variability in each classroom, the proposed paper will report and evaluate the relationships between the criterion (pupil post-test performance) and a wide range of potential variables which might be used for statistical control. These include different forms of the pre-test, several measures of intellectual ability, set toward achievement, etc. Reliability of pupil pre- and post-test performance will also be assessed.

James K. Duncan, John B. Hough, and James H. Thompson, Ohio State University. Exploratory Studies of the Teaching Situation Reaction Test.

The Teaching Situation Reaction Test (TSRT) was developed to help in the study of preservice education courses. The first research report of studies of TSRT was presented at AERA in February 1965. This second report will present studies of the reliability, predictive validity, and item discrimination characteristics of a revised edition of the TSRT.

Thirteen of the 48 items on the test have been revised. Of those revised, five now discriminate between high and low scorers at the .01 level, two at the .05 level, one at the .02 level and two at the .10 level.

The revised and improved TSRT has been administered to more than 300 preservice teachers and to over 100 inservice teachers. One study of the difference of mean scores of good and poor student teachers was significant at the .02 level (N = 50). Data has been gathered for (1) another study of the relationship between TSRT scores and student teaching grades (N = 44); (2) one test-retest reliability study (N = 75); (3) a study of the relationship of scores on TSRT and the Watson Glaser Critical Thinking Appraisal (N = 108); (4) an item discrimination study for all 48 items (N = 300); and (5) a study of the relationship of TSRT scores of inservice teachers to selected aspects of their student's performance (N = 100).

Donald A. Leton and Gilbert Sax, University of Hawaii. Regression and Discriminant Analyses of Entrance Examinations for Students in Elementary and Secondary Education.

The purpose of this study was to evaluate the use of the American Council on Education Psychological Examination (ACE), Ohio State Psychological (OSU), Cooperative General Achievement Test (CGAT), and the National Teachers Examination (NTE) in the teacher-education programs at the University of Hawaii. A longitudinal study of the students entering in the 1957-58 school year, and graduating in the 1960-61 school year, or dropping-out or failing in the intervening years, was carried out. Comparisons were made of the levels of performance by kindergarten-primary, elementary, and secondary majors on the entrance examinations. Stepwise multiple regression analyses were carried out for the prediction of success in the elementary and secondary programs. Criterion variables for prediction analyses included grades in practice and intern teaching courses, and related education seminar grades. Since the NTE examination was administered in the senior year, the Professional Information subtest was also used as a criterion variable. Discriminant analyses was used for the classification of successful vs. unsuccessful students.

Bob Burton Brown, University of Florida. Measurement of Experimentalism

This paper is a description of the development and research use of the Experimentalism Scale. This scale is composed of the (1) Personal Beliefs Inventory, which measures agreement-disagreement with John Dewey's general philosophy, and the (2) Teacher Practices Inventory, which measures agreement-disagreement with John Dewey's philosophy of education.

There will be a brief discussion of Dewey's experimentalism as a theoretical framework for the scale, and a detailed analysis of how items of the inventories fit into the various theoretical classifications of this framework. Developmental procedures will be discussed, including the establishment of content validity, item analysis, reliability for successive forms of the inventories, and the results of submitting the scale to factor analysis.

Several research uses to which the instruments have been put, and their results will be discussed, including my original "Study of the Relationship of Experimentalism to Classroom Practices", several extensions of this study by several of my graduate students, and the proposed use of the Experimentalism Scale as a keystone in my current "Investigation of Observer-Judge Ratings of Teacher Competence.

SESSION 96

Paper session --CHARACTERISTICS OF SOCIALLY DISADVANTAGED CHILDREN

Roberta C.Meyer, Robert D. Hess, and Virginia C. Shipman, University of Chicago. *Social Class Differences in Maternal Attitudes Toward School and the Consequences for Cognitive Development in the Young Child.*

Maternal attitudes toward the school were found to be related to the child's cognitive development. This finding is part of a larger study of maternal influences during the preschool years on educability and cognitive skills. *Ss* were 163 pairs of urban Negro mothers and their four-year olds, divided by sex and by four social status levels. In 3 of the groups, upper-middle, upper-lower, and lower-lower, families were intact (*N* = 40, 42, 40, respectively). In the fourth, also a lower-lower group, fathers were absent and mothers received public assistance (N = 41). A twenty-seven item instrument, constructed to measure attitudes toward education and the public school system, was administered during a home visit. Cognitive measures (WAIS and Binet IQs and Conceptual Style sorting abilities) were obtained a few weeks later when the mother-child pair was tested at the University.

Factor analysis of mothers' responses to the attitude measure produced six factors, including feelings of ineffectiveness in interaction with the school system, respect for formal education and pessimism about its ultimate value. Social class differences appeared in both mothers' attitudes toward the educational process and children's cognitive abilities. Although there were significant relationships between mothers' skills and their children's abilities, more importantly, mothers' attitudes toward the educational process were significant predictors of their children's educability.

Ellis G. Olim, Robert D. Hess, and Virginia C. Shipman, University of Chicago. *The Role of Mothers' Language in Mediating Pre-School Children's Cognitive Development.*

Using a research group of 163 urban Negro mothers and their 4-year old children, drawn from four social class groups, a study was made of the relationship between the mothers' language style (as assessed by various indices of lexical, syntactic, and cognitive complexity and elaboration in three speech samples) and the conceptualizing level of the children (as measured by the Sigel Conceptual Sorting Task for Children, which assesses the child's ability and preference for forming four types of concepts).

It was found: 1) that the mother's language, regardless of speech sample, was the most significant variable related to the children's concept formation; 2) that the child's and mother's IQs were relevant to the child's ability to do the task, but of little or no significance in the case of the child who could form concepts; and 3) that social class level was relevant in the case of only one of the four types of concepts.

On the basis of these findings, the authors develop the thesis that the inadequate level of cognitive development observed in socially disadvantaged children is determined in large measure by a deficiency of cognitive interaction in families of low socio-economic status. The role of language as an important mechanism for mediating between cultural experience and cognitive development is explicated. Finally the implications of the language mediation hypothesis for compensatory program for the culturally disadvantaged child are discussed.

Loren S. Barritt, Melvyn I. Semmel, and Paul Weener, University of Michigan. *A Comparison of the Psycholinguistic Functioning of "Educationally-Deprived" and "Educationally-Advantaged" Children.*

The purpose of this research is to explore the relative psycholinguistic functioning of school-age children in relation to several socio-cultural variables. Three samples of kindergarten-first grade children were drawn from the following school populations: 1) *de facto* segregated school (approx. 75% Negro); 2) integrated school (approx. 50% Negro); 3) *de facto* segregated school (approx. 0% Negro). Performance was determined for all *S*s on the nine subscales of the Illinois Test of Psycholinguistic Abilities (ITPA).

There are significant differences between the 3 groups on 5 of the 9 subscales of the ITPA. Significant differences were observed on the subscales. Intercorrelations of the subscale scores indicates strongest relationships between: 1) Auditory Vocal Automatic and Auditory Vocal Association r = .62; 2) Auditory Vocal Association and Auditory Decoding r = .54; 3) Auditory Vocal Automatic and Auditory Decoding 5 = .50. A comparison of white and Negro pupils' ITPA scores yielded significant differences on seven of the nine subscales.

Daniel C. Neale, John M. Proshek, University of Minnesota. *School-Related Attitudes of Culturally Disadvantaged Elementary School Children.*

A version of the *Semantic Differential* was used to sample attitudes of 350 children in the 4th, 5th, and 6th grades of two elementary schools. Compared to the other schools in the same city, School I was low on socio-economic indicators, School II was near the median.

Children responded on 16 adjective dimensions to each of 15 stimulus phrases, a number of which were school-related. A factor analysis procedure showed a strong evaluative (good-bad) factor, which accounted for one-third of the total variance. No other factor was prominent. Children's ratings on the eight adjective dimensions most strongly associated with the evaluative factor were added to represent each child's "attitude" toward each stimulus. These scores were analyzed for each stimulus in a three-way analysis of variance procedure which tested the effects of school, grade, and sex.

Children in School I had significantly higher evaluative scores for "my school books", "having to keep quiet", "following rules", and "my school building". Children in School II were significantly more positive toward "my teacher", "father", and "college student". As grade in school increased, evaluative scores became significantly less positive for a variety of stimuli, including "my school books", "my classroom", "my teacher", and "me". In addition to expected differences in the attitudes of boys and girls, an interesting interaction between sex and school was noted. Attitudes toward several stimuli were similar for boys and girls in School I but markedly different for boys and girls in School II.

Egon Mermelstein, Hofstra University and Lee S. Shulman, Michigan State University. The Effect of Lack of Formal Schooling on Number Development.

The major objectives of this study were: 1) to examine the effects of a period of non-schooling on attainment of the concept of conservation by Negro children; 2) to investigate differences between verbal and non-verbal assessments of the same cognitive structure; and 3) to assess the differential effects, if any, of systematically varying the types of questions utilized in verbal tests of conservation.

To achieve the first objective, the performances of Negro six and nine-year-old children from Prince Edward County, Virginia, a community which had been without public schools, were compared with those of Negro children from a community which had had regular schooling. As to the second objective, all children were compared on both verbal and non-verbal tests of conservation. To assess the differential effects of questioning, three techniques of questioning were experimentally varied.

Findings revealed, generally, no significant differences attributable to the effects of non-schooling. Differences between verbal and non-verbal tasks were found to be highly significant. One of the three questioning conditions appeared to influence performance on verbal conservation tasks.

SESSION 98

Paper session --STUDIES OF STUDENT APTITUDE AND ACHIEVEMENT IN UNIVERSITY

John M. Newell, Tufts University, Medford, Massachusetts. *Creativity, Predicted and Actual Achievement by High Aptitude College Freshmen.*

The present study is designed to determine if the results reported by Edwards and Tyler hold for a selective, high aptitude group. Such findings would be contrary to those of Getzels and Jackson and would suggest that the inclusion of creativity tests in admissions test batteries may be of little value for selection purposes.

Subjects for the present study were drawn from the freshman class of a small liberal arts college. The median SAT scores for this group fall in the upper 5% on national norms. Two groups, a high creative and a low creative were selected. Each group is composed of 24 subjects, 14 males and 10 females. Selection was based on scores on the Word Association Test and the Uses Test taken from Getzels and Jackson (2). SAT scores, actual grade-point average for the freshman year and a "predicted" grade-point average were obtained. The latter average represents the results of a statistical computation based on high school performance and aptitude test scores which reflects what the subject may be expected to achieve in college.

Results indicate that while the correlation between aptitude and creativity is low, the high creative group had significantly higher SAT scores, had significantly higher actual grade-point averages, and had significantly higher "predicted" grade-point averages. Findings suggest that while the creativity tests and the Scholastic Aptitude Test appear to be sampling different facets of talent (low correlations), both measures are correlated with achievement.

The results are viewed as supporting the findings of Edwards and Tyler (1). Additional analyses when the group is divided into "science" and "non-science" majors are presented. Implications for the use of creativity tests to identify students with specific talents after admission to college are discussed.

Kalil I. Gezi, Chico State College, California and Roger Cummings, San Francisco State College. *Relationships Between Certain Measures of Academic Potential and College Achievement.*

The purpose was to determine the relationship of four measures of academic potential (high school grades, ACT scores, high school and ACT predictors combined, and local predictors) with college achievement (GPA), and to discover the degree of accuracy to which college achievement was predicted for certain student sub-groups.

The findings were: 1) Correlation between high school grades, ACT scores, a combination of both, and the local predictors on the one hand and overall GPA on the other hand were not significantly high. 2) However, the best of the four predictors of college GPA is the combined index of high school grades - ACT scores ($r = 0.526$). 3) The accuracy of the combined index in predicting colleges grades was in the following order: natural sciences, social studies, English and mathematics. 4) The accuracy of the combined index in predicting college grades for student sub-groups was in the following order: academic freshmen, female freshmen, collegiate freshmen, male freshmen, vocational freshmen, male freshmen, female transferers and male transferers. 5) All members of the sample obtained college grades below their predicted academic potential.

Edward A. Nelsen, University of Wisconsin and Eleanor E. Maccoby, Stanford University. The Relationship Between Social Development and Differential Abilities on the Scholastic Aptitude Test

This study was designed to investigate hypotheses centering around sex typing and the interfering effects of certain tension producing experiences upon cognitive processes. A biographical questionnaire concerning family and social relationships was given to 1,956 college freshmen. The responses to this questionnaire were analyzed in relation to differential verbal and mathematical abilities. For boys it was found that a high-verbal, low-mathematics pattern was associated with reports of father absence, punishment exclusively by the mother, fear of father, and reports of having been a "mamma's" or "daddy's boy". A high-mathematics, low-verbal pattern was found among boys who reported punishment exclusively by the father and having talked over personal problems with the father. For girls, a high-verbal, low-mathematics pattern was associated with reports of fearfulness of one's mother and "only sometimes" talking over personal problems with the father. A high-mathematics, low-verbal pattern was found for Ss of both sexes who reported having many friends and some close friends. Consideration of the results led to proposal of a new hypothesis integrating both the sex typing and tension-interference hypotheses.

Clinton I. Chase and Marilyn K. Evans, Indiana University. The University Freshman Dropout.

The present study was carried out in order to characterize first semester dropouts from students who complete that semester; and to follow up the dropouts and characterize those who have returned to college within the 3 year span of this study with those who have not returned.

Of all freshmen students who registered at Indiana University in the fall of 1961, 90 dropped out before the end of the first semester. Of these 90, 75 had complete data on their records. These were compared with the remainder of the class who actually finished the first term. Biographical, educational, and test data were the basis of comparison. Three years later, dropouts were followed up to see how many had returned to college. Returnees were compared with non-returnees. The principal findings were: 1) Mean scores for dropouts were significantly lower than non-dropouts on aptitude and achievement tests, although a fair amount of overlap existed between the two groups, suggesting other variables may be related to dropping out. 2) Seven items of personal and academic history were found to be significantly different between dropouts and non-dropouts. The dropouts appeared to be characterized as relatively weak in achievement and aptitudes that facilitate success in college, and came from a background of minimum involvement in academic affairs. The returnee was not distinguished from the non-returnee.

INDEX OF NAMES

Abrams, A. M. 83, 89
Allen, D. W. 69, 70
Allen, W. H. 105
Ames, W. S. 33
Amidon, E. 13, 87
Anderson, R. C. 43, 98, 107
Anthony, S. M. 80
Armstrong, J. R. 88
Arnoff, M. 79
Ayers, J. D. 73

Baker, E. L. 122
Baker, F. B. 53
Baker, R. L. 25, 82, 116
Baldwin, T. L. 100
Barnhard, H. J. 46
Barritt, L. 82, 112, 125
Beaird, J. H. 30
Beker, J. 75
Beller, E. K. 12, 13
Bennett, C. M. 51
Bennett, L. M. 28
Bennett, S. W. 82
Berger, R. J. 25
Bergan, J. R. 2
Biaggio, A. B. 89
Bisbicos, E. E. 108
Bolvin, J. O. 10
Bolyard, A. J. 103
Bormuth, J. R. 74
Boyan, N. J. 72
Brookover, W. B. 50
Brown, B. B. 123
Brown, G. I. 71
Brumbaugh, R. B. 76
Buelke, J. A. 101
Burnett, J. H. 78
Bussis, A. M. 85
Byers, J. L. 45

Campbell, J. T. 85
Carton, A. S. 86
Cartwright, G. P. 112
Casey, J. P. 58
Centra, J. A. 15
Chase, C. I. 56, 128
Chasnoff, R. E. 32
Chen, M. K. 39
Childs, J. W. 77
Cohen, J. 90
Collier, R. O. 53
Corman, B. R. 35
Cowles, M. 13, 79
Cox, R. 10
Creel, J. 80
Cresswell, W. H., Jr. 114
Crispin, D. 101
Cronin, J. M. 17

Daniel, K. B. 79
Davidson, R. E. 34, 45
Davis, O. L., Jr. 3, 62, 117
Dell, D. L. 51
Denny, T. 96
Deno, S. L. 99
Deutsch, C. P. 2
DeVault, M. V. 88
Duell, O. K. 107
Duffy, E. J. 7
Duncan, J. K. 122
Dunn-Rankin, P. 53

Eash, M. J. 51
Eisner, E. W. 85
Ellinger, B. D. 31
Elliott, C. B. 117
Erickson, D. A. 76
Erickson, E. L. 50
Espenschied, J. R. 22
Evans, M. K. 128

Faix, T. L. 38
Faust, G. W. 98
Feldmann, S. C. 2, 89
Fenske, R. H. 51
Ferris, M. J. 74
Filep, R. T. 105
Finger, J. A. 64
Fogarty, B. M. 6
Fortune, J. C. 100
Francke, D. C. 6
Franzen, W. L. 120
Friedman, H. L. 106
Fry, E. B. 47

Frymier, J.R. 67

Gage, N.L. 100
Gerlach, V.S. 47, 116
Gezi, K.I. 127
Gifford, E.M. 95
Glaser, R. 9
Glass, G. 4
Glick, O. 87
Gnagey, W.J. 5
Goodkind, T.B. 38
Goodwin, W.L. 56
Gordon, J.M. 103, 116
Gordon, R.E. 118
Greene, J.E., Sr. 16
Griffin, W.J. 81
Grobman, H. 46
Guthrie, J.T. 43
Guthrie, J.W. 18
Gutsch, K.U. 67

Harris, A.J. 90, 113
Hastings, J.T. 114
Hayes, R.B. 107
Heard, H. 47
Henderson, E.H. 65
Herbert, J. 88
Helmick, J.S. 85
Herriott, R.E. 75
Herzog, B.G. 81
Hess, R.D. 124
Hodges, R.E. 41
Hough, J.B. 42, 122
Hoy, W.K. 36
Huffman, W.J. 114
Husek, T.R. 35, 66
Hutchins, E.B. 83

Jabker, E.H. 27
Jackson, J. 80
Jackson, P.W. 119
Jenkins, J.R. 99
Johnson, M.C. 48
Johnson, P.E. 99
Johnson, T.J. 100
Joiner, L.M. 50
Josephine, Sister 12
Joyce, B.R. 41

Karraker, R.J. 104
Kay, C. 79
Kean, J.M. 62
Keislar, E.R. 31, 58
Kenney, J.B. 72
Keppel, G. 97
Kerlinger, F.N. 121
Kersh, B.Y. 68
Kight, H.R. 119
King, A.R. 37
Kline, C.E. 8
Klopfer, L.E. 91
Knirk, F.G. 103
Kowitz, N.G. 63
Kowitz, G.T. 63
Krugh, C.A. 50
Krumboltz, J.D. 29

Lahaderne, H.M. 119
Lang, G. 56
LaRocque, G.E. 20
Larson, E. 88
LaShier, W.S., Jr. 61
Lasswell, A. 96
Lehmann, I.J. 14
Leton, D.A. 123
Levin, J.R. 65
Levonian, E. 26
Lewis, E.C. 14
Lewy, A. 66
Lindeman, R.H. 54, 118
Lindvall, C.M. 9
Lipson, J.I. 11
Little, J.K. 52
Long, B.H. 65
Long, J.M. 46
Lovinger, R.J. 90
Lundsteen, S.W. 59
Lynch, S. 34

Maccoby, E.E. 128
Macdonald, J.B. 115
Marascuilo, L.A. 65, 93
Marco, G.L. 57
Maxey, E.J. 22
McCandless, B.R. 63
McCarty, D.J. 18
McColly, W. 20
McDonald, F.J. 69, 70
McGuire, C. 66
McLean, L.D. 93
McNeil, J.D. 31, 117
Mehrens, W.A. 40

Melton, R.S. 85
Mermelstein, E. 126
Merrifield, P.R. 3, 22, 117
Merrill, M.D. 110
Meyer, R.C. 124
Miller, E.L. 19
Miller, H. 114
Millman, J. 25
Mitchell, J.V. 84
Molineux, J.E. 48
Mountain, L. 47
Mowrer, D.E. 82
Muller, D.H. 33
Myler, H. 80

Neal, C.M. 95
Neale, D.C. 125
Nelsen, E.A. 128
Newell, J.M. 118, 127
Nonneman, A.J. 83

Ober, R. 42
Ohnmacht, F.W. 44
Ojemann, R.H. 22
Olim, E.G. 124
Olson, A.V. 1
Orme, M.E.J. 69, 70
Orr, D.B. 106
Otto, W. 26
Overing, R.L.R. 108
Owens, R.E. 60

Payne, D.A. 92
Peccolo, C.M. 71
Peckens, R.G. 28
Perfetti, C.A. 82
Personke, C.R. 19
Pfeiffer, I.L. 115
Phillips, J.A., Jr. 3
Popham, W.J. 42, 122
Porter, A.C. 54
Porter, D. 111
Powell, W.R. 111
Prentice, J.L. 34
Proshek, J.M. 125
Pruzek, R.M. 73

Rahmlow, H.F. 49
Ramsey, J.G. 44
Raven, R.J. 23
Regal, J.M. 90
Rehberg, R.A. 50
Remstead, R. 20
Reynolds, J.A. 77
Rezler, A.G. 28
Richards, J.M., Jr. 39
Rippey, R.M. 21
Rizzo, M. 118
Robinson, G.E. 121
Rogers, B. 48
Rose, H.B. 64
Rossmiller, R.A. 17
Rothkopf, E.Z. 108
Rowe, M.B. 23
Ryan, T.B. 109

Samph, T. 87
Sassenrath, J.M. 97, 119
Sax, G. 123
Scandura, J.M. 109
Schmeidler, R.J. 54
Schmidt, D.E. 2
Schutz, R.E. 25, 82, 116
Schurdak, J.J. 49
Scott, R.E. 121
Scott, W.O. 114
Scribner, J.D. 17
Semmel, M.I. 81, 82, 112, 125
Serwer, B.L. 113
Sheppard, L.E. 29
Shipman, V.C. 124
Shuell, T.J. 97
Shulman, L.S. 126
Shutes, R.E. 100
Simon, A. 87
Sliepcevich, E.M. 37
Slobodian, J.J. 62
Smith, M.D. 103
Smith, N.B. 31
Snider, B.C.F. 22
Snider, R.M. 55
Soar, R.S. 61, 87
Spaights, E. 29
Sproull, N. 50
Staffieri, J.R. 63
Stake, R. 25
Stanley, J.C. 89
Stern, C. 58
Stickell, D.W. 91
Stolurow, L.M. 110
Strandberg, J. 31

Sullivan, H. J. 47, 116

Tanaka, M. 85
Tanner, D. 102
Thompson, J. H. 122
Thurner, R. 93
Tornow, E. W. 7
Tovatt, A. 19
Travers, R. M. W. 108
Tremonti, J. B. 1
Turner, R. L. 4
Turpan, M. 56
Twelker, P. A. 68
Vaughn, H. A. 92
Vinsonhaler, J. F. 48

Walberg, H. J. 15
Wallin, H. A. 78
Weber, W. 13
Weener, P. 112, 125
Weintraub, S. 96
Wehling, L. J. 36
Weiner, M. 89
West, L. J. 106
White, W. F. 24
Wightman, L. 55
Wilcoxon, F. 53
Wittrock, M. C. 43
Wodtke, K. H. 40
Wolfe, W. 31
Wolins, L. 14

Yeager, J. 9
Yee, A. H. 3
Yelsma, J. J. 14
Yonge, G. D. 59

Ziller, R. C. 65